



DIEM
GMA

Quaderni di ricerca

Quarta giornata di studio

Ettore Funaioli

16 luglio 2010

A cura di
Umberto Meneghetti, Alberto Maggiore e Vincenzo Parenti Castelli

Quaderni del **DIEM – GMA**

Atti di giornate di studio – 4

A cura di:

U. Meneghetti , A. Maggiore , V. Parenti Castelli

Coordinatore di redazione:

Alessandro Zanarini

DIEM

Dipartimento di Ingegneria delle
Costruzioni Meccaniche, Nucleari, Aeronautiche e di
Metallurgia

www.diem.unibo.it

GMA

Gruppo di Meccanica Applicata

http://wpage.unina.it/dellaval/GMA/GMA_home.htm

Quarta giornata di studio Ettore Funaioli

16 luglio 2010

*A cura di:
Umberto Meneghetti, Alberto Maggiore e Vincenzo Parenti Castelli*



Proprietà letteraria riservata.
© Copyright 2011 degli autori
Tutti i diritti riservati

Quarta giornata di studio Ettore Funaioli - 16 luglio 2010
A cura di Umberto Meneghetti, Alberto Maggiore e Vincenzo Parenti Castelli
Bologna : 2011 - p. 456; 17 cm.

ISBN 978-88-965720-6-1

Versione elettronica disponibile alla pagina
<http://amsacta.cib.unibo.it/3043/>

Stampa a richiesta eseguita da:
Asterisco Snc Tipografia Digitale
Via Belle Arti, 31 a/b – 40126 Bologna
Tel 051 236866 – Fax 051 261105
mail: grafica@asteriscosnc.it
www.asteriscosnc.it

INDICE

Prefazione

<i>A.Freddi</i>	1
Ricordo del Prof. Gustavo Favretti a diciannove anni dalla morte	
<i>G. Catania, S. Sorrentino</i>	3
A dynamic model of railway bridges crossed by travelling trains	
<i>A. Zanarini, M. Neumann, J. Wassermann</i>	17
Full-field optical measurements for advanced structural dynamics: first outcomes	
<i>M. Cavallari, A. Montanari, E. Mucchi, G. Dalpiaz</i>	33
Analysis of the evolution of the pressure forces in variable displacement vane pumps using different approaches	
<i>R. Vertechy, A. Frisoli, M. Solazzi, A. Dettori, M. Bergamasco</i>	57
Design and control of new robotic joints for rehabilitation exoskeletons	
<i>M. Troncossi, A.Rivola</i>	73
Experimental tests of joint prototypes for upper limb powered prostheses	
<i>N. Sancisi, D. Zannoli, V. Parenti Castelli</i>	83
Sensitivity analysis of 1-dof equivalent mechanisms for the kinematic modelling of the human knee	
<i>U. Meneghetti</i>	97
Applicazioni storiche delle camme	
<i>M. Mozaffari Fomashi, M. Troncossi, V. Parenti Castelli</i>	113
A study on the rehabilitation hand exoskeletons	
<i>A. Chebbi, V. Parenti Castelli</i>	127
Analysis of the internal torque and bending moment acting on the legs of the 3-UPU translational parallel manipulator	
<i>C. Innocenti</i>	139
Synthesis of the spherical four-bar linkage function generator for six precision points	
<i>E. Prati, M Silvestri, A. Tasora, T. Marin</i>	155
Influenza dell'attrito sullo stato di tensione e di deformazione in anelli di tenuta in elastomero	

<i>A.O. Andrisano, M. Faretra, F. Gherardini, A. Guerra, F. Leali, M. Pellicciari, F. Pini, A. Vergnano</i>	165
Metodi di progettazione integrata di sistemi automatici adattivi	
<i>A. Rossi, G. Rosati, S. Cenci, A. Carli, L. Zanotti, M. Mantovani, A. Foroni, V. G. Riello</i>	181
La mecatronica per l'assemblaggio flessibile	
<i>E. Dragoni, W. J. Bagaria</i>	193
Mechanical design of bimaterial helical springs with circular cross-section	
<i>G. Medri</i>	205
Considerazioni sui criteri di resistenza	
<i>G. Nicoletto</i>	219
Progettazione a fatica termomeccanica di getti in lega d'alluminio per applicazioni motoristiche	
<i>A. Strozzi, L. Bertocchi, E. Campioni</i>	231
Normalization of the stress concentrations at the interference fit between a cylindrical shaft and a hub with rounded edges	
<i>A. Freddi</i>	241
Sperimentare per progettare	
<i>M. Barbieri, F. Pellicano, A. O. Andrisano</i>	259
Influence of dynamic effects on point EHL contacts in gear systems	
<i>G. Scirè Mammano, E. Dragoni</i>	265
Fatica termomeccanica di fili a memoria di forma: apparato di prova e primi risultati	
<i>D. Castagnetti, A. Spaggiari, E. Dragoni</i>	283
Effect of bondline thickness on the static strength of structural adhesives under nearly-homogeneous shear stresses	
<i>G. Reggiani, M. Cocconcelli, R. Rubini, C. Borghi, A. Ferrari</i>	303
Use of neural networks in cerebral palsied children: form recognition by means of gait analysis	
<i>G. Scirè Mammano, M. Cocconcelli, R. Rubini, E. Dragoni</i>	315
Accelerometri MEMS: caratterizzazione dinamica e confronto con i sensori piezoelettrici	
<i>A. Segatori, B. Reggiani, L. Donati, F. Bagagli, L. Tomesani</i>	337
Investigation and prediction of charge welds in aluminum extruded profiles	

<i>A. Carminelli, G. Catania</i>	359
Aggiornamento automatico di un modello ad elementi finiti B-spline mediante rilievi sperimentali di funzioni di risposta in frequenza	
<i>I. Todaro, R. Squatrito, L. Ceschini, A. Morri, L. Tomesani</i>	373
Effetto del contenuto di idrogeno e delle velocità di raffreddamento sulla microporosità nella produzione per fusione di campioni a difettologia controllata per prove di fatica	
<i>G. Olmi, V. Dal Re</i>	385
Analisi teorica e sperimentale delle sollecitazioni in esercizio su un anello rotante di una macchina cellofanatrice	
<i>D. Croccolo, M. De Agostinis, N. Vincenzi</i>	411
On shaft-hub hybrid couplings	
<i>G. Berselli, M. Piccinini, G. Vassura</i>	429
Valutazione della cedevolezza selettiva in giunti elastici per applicazioni robotiche	
<i>M. Carricato</i>	443
Under-constrained cable-driven parallel robots	
Indice degli autori	455

Prefazione

Questo volume raccoglie le memorie presentate nel corso della Quarta Giornata di Studio Ettore Funaioli, svoltasi presso la Facoltà di Ingegneria di Bologna il 16 luglio 2010. Dalla Prima alla Quarta Giornata di Studio Ettore Funaioli si è passati da 13 a 31 memorie. La crescente adesione di amici ed allievi – di prima e di seconda generazione – alle Giornate di Studio intitolate a Ettore Funaioli, conferma l'alta considerazione per la sua figura sotto l'aspetto umano e scientifico, e conferma altresì la diffusa volontà di ritrovarsi fra colleghi ed amici che si stimano e si apprezzano, per fare tutti partecipi dei risultati scientifici conseguiti e per discutere insieme dei comuni problemi che incontriamo nella nostra attività quotidiana.

A tutti i partecipanti alla Quarta Giornata di Studio Ettore Funaioli va il nostro più cordiale ringraziamento: a loro siamo debitori per il successo della manifestazione.

È motivo di grande soddisfazione constatare ancora una volta il costante impegno dei Ricercatori e l'elevata qualità dei lavori presentati, testimonianza evidente della vitalità delle Discipline che riconoscono nella Meccanica la loro radice comune. È questo un buon auspicio anche per il nostro Paese, che nell'Industria meccanica ha una delle principali fonti di benessere e di primato.

È da ricordare che la Giornata ha visto anche, alla presenza dei familiari, la commemorazione del compianto Prof. Gustavo Favretti nel diciannovesimo anniversario della scomparsa.

La Giornata ha potuto svolgersi grazie anche alla collaborazione della Facoltà di Ingegneria dell'Alma Mater Studiorum Università di Bologna e del DIEM - Dipartimento di Ingegneria delle Costruzioni meccaniche, nucleari, aeronautiche e di Metallurgia. Ringraziamo vivamente il Preside della Facoltà, Prof. Pier Paolo Diotalle, e il Direttore del DIEM, Prof. Gianni Caligiana, che hanno consentito queste collaborazioni ed hanno voluto aprire la Giornata porgendo il loro saluto ai partecipanti.

Desideriamo infine ringraziare l'Ing. Alessandro Zanarini che ha cooperato nel lavoro di organizzazione delle memorie raccolte in questo volume.

Bologna, 10 giugno 2011

Umberto Meneghetti – Alberto Maggiore – Vincenzo Parenti Castelli

RICORDO DEL PROF. GUSTAVO FAVRETTI A DICIANNOVE ANNI DALLA MORTE

Intervento del Prof. Alessandro Freddi

Alla contemporanea conclusione dell'attività accademica di tre colleghi e amici del Prof. Gustavo Favretti e a diciannove anni dalla Sua scomparsa è mio desiderio porgere questo omaggio e ricordarLo in questa giornata, anche se la Sua persona rimarrà viva nella nostra memoria.

Quando i colleghi Umberto Meneghetti, Sergio Curioni e il sottoscritto entrarono negli stessi anni nell'Istituto di Meccanica Applicata alle Macchine, i professori Ettore Funaioli, Fausto Caboni e Gustavo Favretti erano i rappresentanti ufficiali dell'istituzione. L'atmosfera che vi si respirava era espressione della loro personalità: tre gentiluomini che, con rispetto reciproco, avevano creato il livello e la stima dell'Istituto e godevano di grande considerazione tra i colleghi e tra gli studenti. Del Prof. Funaioli in più occasioni sono state sottolineate le doti di scienza e di guida; il Prof. Caboni si distingueva per un impegno didattico eccezionale e il Prof. Favretti era in particolare impegnato a creare le basi della ricerca nella Costruzione di Macchine.



Gustavo Favretti 1927 – 1991

Il Prof. Favretti fu da me commemorato per le sue doti morali e professionali a Palermo nel settembre 1991, a Forlì nel 1993 e dai suoi ex allievi, (allora già Professori Ordinari a Modena - Reggio e Parma), in occasione del Convegno AIAS di Parma. Oggi vorrei ricordarLo per alcune Sue specifiche doti umane.

Già a Forlì, il Rettore dell'Università di Bologna di allora Prof. Fabio Roversi Monaco, ebbe cura di mettere in luce, tra le altre, una Sua dote caratteristica ben nota: avere operato per un ambiente di lavoro amichevole tra tutti i collaboratori; aspetto non marginale di senso dell'istituzione.

Ricordo in particolare, con molta simpatia, l'invito che Egli rivolse a me e al Prof. Meneghetti il primo anno del mio contatto con l'Istituto, per un viaggio estivo in Francia, viaggio che si sarebbe rivelato di formazione e di fondazione di un'amicizia. Continuummo poi per tutti gli anni a venire a consolidarla, con l'aiuto prezioso della Signora Rema, che nel frattempo era entrata nella Sua vita con forte comunione di spirito e condivisione d'idee, di lì a poco Ordinario di Glottologia e Linguistica.

Tra i tanti aspetti, voglio ricordare in particolare, un tratto caratteristico della personalità del Prof. Favretti che sintetizzo in questo modo: egli affrontava la realtà sempre con atteggiamento di persona colta. Aveva per esempio una grande cultura musicale, ben oltre la passione del dilettante e sosteneva, con la franchezza che lo caratterizzava, che non era interessato alla musica antica ma che "si limitava" ad ascoltare la musica da Bach ai contemporanei: al pari di un professore di Storia Moderna che, aldilà della sua cultura personale, ha il compito istituzionale di approfondire il periodo di sua competenza. E quel periodo era davvero approfondito. Era un esercizio di serietà intellettuale che non limitava alla Sua disciplina, nella quale si distingueva peraltro per professionalità e "integrità scientifica" (nel senso di una famosa prolusione di Richard Feynmann), ma che estendeva con innata modestia e riservatezza a ogni aspetto della realtà.

Nello stesso modo, ogni Suo compito veniva svolto con precisione e tempestività. Le relazioni ai Convegni preparate con cura, i contributi altrui seguiti tutti con attenzione e interesse critico, così all'estero come ai Convegni nazionali AIAS e AIMETA, associazioni scientifiche delle quali era stato co-fondatore.

Dopo il Convegno, il tempo era dedicato a riflessioni generali se non allo studio (come in Germania per seguire un Corso), a un concerto o a un'opera su cui si era documentato per il luogo e il tempo; tutti questi aspetti erano da Lui intesi come elementi di naturale formazione per uno studioso.

Sono queste alcune caratteristiche del Suo spirito che voglio, assieme ai colleghi, ricordare e testimoniare, specialmente a chi non ha avuto la fortuna di conoscerLo e di apprezzarLo.

A DYNAMIC MODEL OF RAILWAY BRIDGES CROSSED BY TRAVELLING TRAINS

Giuseppe Catania

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: giuseppe.catania@unibo.it

Silvio Sorrentino

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: silvio.sorrentino@unibo.it

Abstract. *This study investigates the dynamic behaviour of railway bridges crossed by travelling trains. A simplified formulation was adopted in order to perform a direct analysis of the effects of the parameters involved in the problem. The bridge is modelled as a rectangular plate, while the trains are modelled as travelling inertial distributed loads. The formulation is accomplished by the use of the Rayleigh-Ritz method, yielding a low order model with time-dependent coefficients. Numerical examples are presented and discussed, aimed at investigating the effects of each of the model-governing parameters.*

Keywords: *railway bridges, moving loads, Rayleigh-Ritz method.*

1. INTRODUCTION

In the analysis of the dynamic effects of railway vehicles on bridges, simplified models are suggested and usually used, taking into account only certain aspects, such as deterministic, vertical effects, and the influence of moving forces and masses [1-2]. Inertial effects of both bridge and vehicle can be influential, and not negligible, since the mass of the external load introduces a coupling effect between the load and the structure. Other important aspects, such as dynamic properties of travelling vehicles and track irregularities are not considered in the present study.

The railway bridge model most commonly used is a continuous Euler Bernoulli beam [2], or a Timoshenko beam [3], traversed by either concentrated [4] or distributed moving loads [5]. Possible applications of lumped vibration absorbers have also been investigated [6]. However in the present study a homogeneous Kirchhoff plate is considered, in order to allow the analysis of lateral vibrations due to trains travelling on double-track bridges. Structure damping is included in the model, as it may play an important role.

The train is simply modelled by means of a continuous load in the form of a moving strip, an idealization which can be adopted when the span of the bridge is large in comparison with the distance between axles [2]. Constant speed of motion of vehicles along the bridge is assumed.

The formulation is accomplished by the use of the Rayleigh-Ritz method [7], and the solution is expressed in terms of a linear combination of functions, which in the present

study are selected as tensor products of eigenfunctions of prismatic pinned-pinned and free-free beams in flexural vibration. This approach yields a reduced order model with time-dependent coefficients, allowing a parametric analysis of plates loaded by travelling distributed masses [8].

Different example cases are presented and discussed in detail, analyzing the effects of velocity, mass and length of the train on the plate's dynamic response with respect to the mass, stiffness and damping of the plate itself.

2. THEORETICAL MODEL

A homogeneous isotropic Kirchhoff plate is considered, simply supported on two opposite sides, free on the other two sides and crossed by a travelling distributed load. The load per unit area p over the plate may be expressed as:

$$p(x, y, t) = -\rho_l f(\xi, y) \left[g + \frac{d^2 w(x, y, t)}{dt^2} \right], \quad \xi = x - v_x t \quad (1)$$

where w is the vertical displacement of a point of the plate or of the load, ρ_l is the equivalent mass per unit area of the load, g is the gravity acceleration, v_x is the travelling speed in the x direction, ξ is a moving coordinate in the same direction [1] and f models the translating strip representing the instantaneous position of the load:

$$f(\xi, y) = [u(\xi + L_t) - u(\xi)] [u(y - \delta) - u(y - \delta - l_t)] \quad (2)$$

Note that within the present study ρ_l is assumed to be constant; however, piecewise-constant or other distributions $\rho_l(\xi)$ may be considered and adopted in the following developments. Equation 2 contains the unit step distribution $u(\cdot)$, L_t and l_t are the length and the width of the strip modelling the train, and δ is the distance between the side of the strip and the edge $y = 0$ of the plate, as shown in Fig. (1). The second term on the right-hand side of Eqn. (1) describes the inertial action of the load. The total acceleration may be expressed in the following general form:

$$\frac{d^2 w}{dt^2} = \frac{\partial^2 w}{\partial t^2} + v_x^2 \frac{\partial^2 w}{\partial x^2} + 2v_x v_y \frac{\partial^2 w}{\partial x \partial y} + v_y^2 \frac{\partial^2 w}{\partial y^2} + 2v_x \frac{\partial^2 w}{\partial x \partial t} + 2v_y \frac{\partial^2 w}{\partial y \partial t} + a_x \frac{\partial w}{\partial x} + a_y \frac{\partial w}{\partial y} \quad (3)$$

where v_x , v_y , a_x , a_y express the velocities and accelerations of the travelling load in the x and y directions respectively [1]. Considering a train travelling at constant speed v in the x direction, Eqn. (3) reduces to:

$$\frac{d^2 w}{dt^2} = \frac{\partial^2 w}{\partial t^2} + 2v \frac{\partial^2 w}{\partial x \partial t} + v^2 \frac{\partial^2 w}{\partial x^2} \quad (4)$$

The first term of the right-hand side of Eqn. (4) expresses the influence of vertical acceleration of the moving load, the second term the influence of Coriolis acceleration, and the third term the influence of track curvature [1].

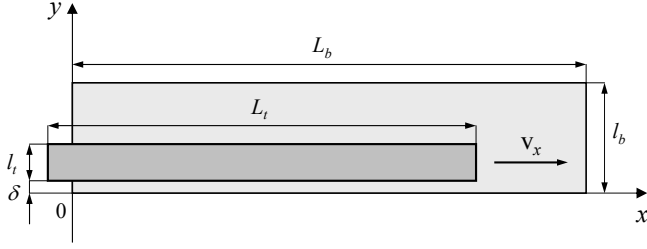


Figure 1. Model scheme.

The functional of the total potential energy of the coupled system can be written as the sum of a term U due to the strain energy plus a term V representing the potential of all applied loads (including the inertial forces):

$$\Pi = U + V \quad (5)$$

The potential of the strain energy can be written in terms of second order derivatives of the out-of-plane displacement w :

$$U = \frac{1}{2} \int_0^{l_b} \int_0^{L_b} D \left[w_{xx}^2 + w_{yy}^2 + 2\nu w_{xx} w_{yy} + 2(1-\nu) w_{xy}^2 \right] dx dy \quad D = \frac{Eh^3}{12(1-\nu^2)} \quad (6)$$

where the subscripts denote differentiation with respect to the spatial variables and D is the flexural stiffness of the plate, expressed as a function of Young's modulus E , Poisson's ratio ν and thickness h [7]. In the adopted formulation the inertial forces are included in the potential of applied loads V as follows:

$$V = \int_0^{l_b} \int_0^{L_b} (\rho_b w \ddot{w} - wp) dx dy \quad (7)$$

where ρ_b is the mass per unit area of the plate and p is the load in Eqn. (1). The out-of-plane displacement w is expressed by means of a linear combination of shape functions, selected as products of homogeneous uniform prismatic beam eigenfunctions ϕ :

$$w = \sum_{n=1}^N q_n \phi_n(\xi, \eta) \Rightarrow w = \boldsymbol{\phi}^T \mathbf{q} \quad (8)$$

where \mathbf{q} is the generalized coordinate vector. Introducing the displacement expansion in the quadratic functional Π , and imposing its stationarity, yields the following algebraic eigenproblem:

$$[\mathbf{M} + r\Delta\mathbf{M}]\ddot{\mathbf{q}} + 2r\alpha\beta[\Delta\mathbf{C}]\dot{\mathbf{q}} + \beta^2[\mathbf{K} + r\alpha^2\Delta\mathbf{K}]\mathbf{q} = -rg\mathbf{f}, \quad r = \frac{\rho_t}{\rho_b}, \alpha = \frac{\nu}{L_b\beta}, \beta = \frac{1}{L_b^2} \sqrt{\frac{D}{\rho_b}} \quad (9)$$

In Eqn. (9) β is a frequency parameter and α is a dimensionless parameter depending on the speed v . The matrices in square brackets can be regarded as dimensionless quantities, and they can be computed according to the following integrals:

$$\begin{aligned}\mathbf{M} &= \int_0^{l_b} \int_0^{L_b} (\boldsymbol{\Phi} \boldsymbol{\Phi}^T) dx dy, \quad \Delta \mathbf{M} = \int_0^{l_b} \int_{x_0}^{x_1} (\boldsymbol{\Phi} \boldsymbol{\Phi}^T) dx dy, \quad \Delta \mathbf{C} = L_b \int_0^{l_b} \int_{x_0}^{x_1} (\boldsymbol{\Phi} \boldsymbol{\Phi}_x^T) dx dy \\ \mathbf{K} &= L_b^4 \int_0^{l_b} \int_0^{L_b} [\boldsymbol{\Phi}_{xx} \boldsymbol{\Phi}_{xx}^T + \boldsymbol{\Phi}_{yy} \boldsymbol{\Phi}_{yy}^T + \nu (\boldsymbol{\Phi}_{xx} \boldsymbol{\Phi}_{yy}^T + \boldsymbol{\Phi}_{yy} \boldsymbol{\Phi}_{xx}^T) + 2(1-\nu)(\boldsymbol{\Phi}_{xy} \boldsymbol{\Phi}_{xy}^T)] dx dy \\ \Delta \mathbf{K} &= 2L_b^2 \int_0^{l_b} \int_{x_0}^{x_1} (\boldsymbol{\Phi} \boldsymbol{\Phi}_{xx}^T) dx dy, \quad \mathbf{f} = \int_0^{l_b} \int_{x_0}^{x_1} \boldsymbol{\Phi} dx dy\end{aligned}\quad (10)$$

where the integration interval $[x_0, x_1]$ is time-dependent. Introducing the ratio between the lengths L_t and L_b :

$$\mu = L_t L_b^{-1} \quad (11)$$

then x_0 and x_1 vary according to Tab. (1).

To model energy dissipation within the structure, a dimensionless damping matrix \mathbf{C} may be defined by means of the plate modal matrix $\boldsymbol{\Phi}$ (mass normalized) and eigenvalues ω_n (computed from the \mathbf{M} and \mathbf{K} matrices), and considering a modal damping ratio ζ equal for all modes:

$$2\zeta \mathbf{C} = 2\zeta \boldsymbol{\Phi}^{-T} \text{diag}(\omega_n \beta^{-1}) \boldsymbol{\Phi}^{-1} = \boldsymbol{\Phi}^{-T} \text{diag}(2\zeta \omega_n \beta^{-1}) \boldsymbol{\Phi}^{-1} \quad (12)$$

Introducing Eqn. (12) in Eqn. (9) yields:

$$[\mathbf{M} + r\Delta \mathbf{M}] \ddot{\mathbf{q}} + 2\beta[\zeta \mathbf{C} + r\alpha \Delta \mathbf{C}] \dot{\mathbf{q}} + \beta^2[\mathbf{K} + r\alpha^2 \Delta \mathbf{K}] \mathbf{q} = -r\mathbf{g}\mathbf{f} \quad (13)$$

Equation (13) is a reduced order discretized model with time-dependent coefficients, which can be solved numerically.

Table 1. Integration time-dependent interval.

case $\mu < 1$	case $\mu = 1$	case $\mu > 1$
$0 \leq vt < L_t \Rightarrow \begin{cases} x_1 = vt \\ x_0 = 0 \end{cases}$	$0 \leq vt < L_b \Rightarrow \begin{cases} x_1 = vt \\ x_0 = 0 \end{cases}$	$0 \leq vt < L_b \Rightarrow \begin{cases} x_1 = vt \\ x_0 = 0 \end{cases}$
$L_t \leq vt < L_b \Rightarrow \begin{cases} x_1 = vt \\ x_0 = vt - L_t \end{cases}$	$L_b \leq vt < 2L_b \Rightarrow \begin{cases} x_1 = L_b \\ x_0 = vt - L_b \end{cases}$	$L_b \leq vt < L_t \Rightarrow \begin{cases} x_1 = L_b \\ x_0 = 0 \end{cases}$
$L_b \leq vt < L_b + L_t \Rightarrow \begin{cases} x_1 = L_b \\ x_0 = vt - L_t \end{cases}$		$L_t \leq vt < L_t + L_b \Rightarrow \begin{cases} x_1 = L_b \\ x_0 = vt - L_t \end{cases}$

3. NUMERICAL RESULTS

Some numerical examples are presented for studying the dynamic behaviour of the model described in section 2. The influence of parameters v , r , μ , β , ζ governing Eqn. (13) is highlighted by studying time responses $w(t)$ and dynamic response functions H of the dimensionless frequency α (playing the role of ‘frequency response functions’) defined according to:

$$H(x, y; \alpha) = \frac{\max_t [w(t)]}{w_s} \Big|_{x, y; \alpha} \quad (14)$$

where w_s is the static deflection due to the load centered in $L_b/2$ [8].

Numerical solutions of Eqn. (13) are computed by means of a Runge-Kutta algorithm, expanding the solution w of Eqn. (8) with 4×2 beam eigenfunctions (4 pinned – pinned eigenfunctions along the x direction and 2 free – free eigenfunctions along the y direction).

Realistic values for parameter β are computed by means of the empirical expression:

$$\beta = \frac{2\pi a}{\lambda_1^2} L_b^{-\gamma} \quad (15)$$

based on large collections of experimental data [2], where a and γ are parameters depending on the kind of bridge considered, as reported in Tab. (2). The values (in Hz) of the first natural frequency and of parameter β for different kinds of bridges are reported as functions of the length L_b in Figs. (2) and (3) respectively.

Effect of the speed of the load

As a reference case study, the following values for the parameters are assumed:

Plate: $L_b = 50$ m, $l_b = 10$ m, $\beta = \sqrt{5}$ rad/s, $\zeta = 0.05$.

Moving load: $\mu = 1.4$, $r = 0.5$, $l_t = 2.5$ m, $\delta = 1.5$ m.

Time responses $w(t)$ are computed at coordinate $x = L_b/2$, $y = l_b/2$ with speed v varying from 30 m/s to 50 m/s (108 Km/h to 180 Km/h), as shown in Fig. (4).

Table 2. Parameters in Eqn. (15), as reported in [2].

Kind of bridge	a	γ
General bridges (average case)	133	0.9
Steel truss bridges	307	1.1
Steel plate girder bridges with ballast	59	0.7
Steel plate girder bridges without ballast	208	1
Concrete bridges with ballast	190	1.1
Concrete bridges without ballast	225	1.2

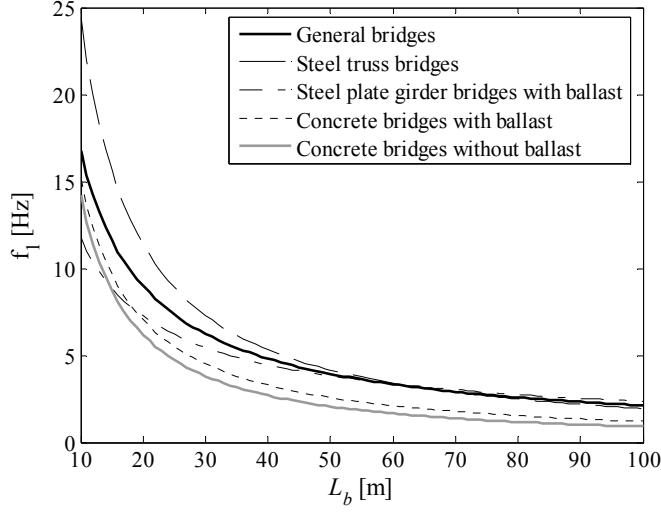


Figure 2. First natural frequency f_1 [Hz] vs. length L_b for different kinds of bridges.

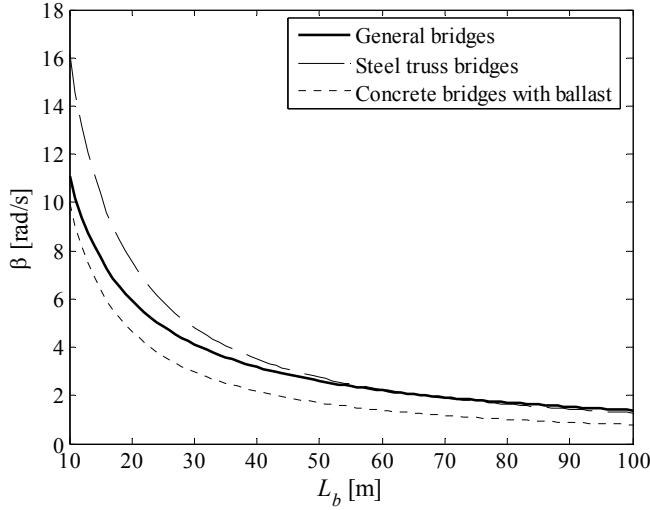


Figure 3. Frequency parameter β [rad/s] vs. length L_b for different kinds of bridges.

Effect of the mass of the load

Parameter values are assumed as in the reference case with $v = 40$ m/s = 144 Km/h, varying r from 0.1 to 1. Response functions $w(t)$ and $H(\alpha)$ are computed in $x = L_b/2$, $y = l_b/2$, as reported in Figs. (5) and (6) respectively. Functions $H(x, y; \alpha)$ show a peculiar undulating trend, which was found to be not significantly affected by the choice of coordinate x [8].

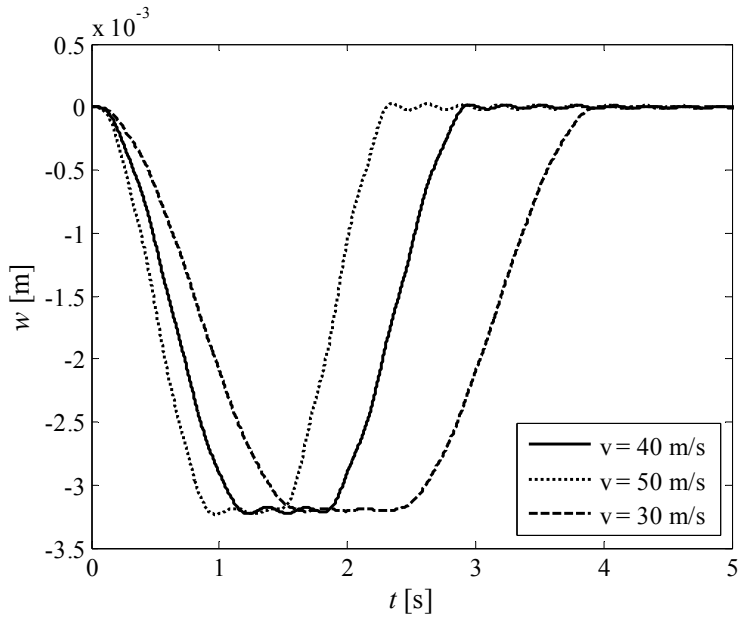


Figure 4. Effect of parameter v on $w(t)$.

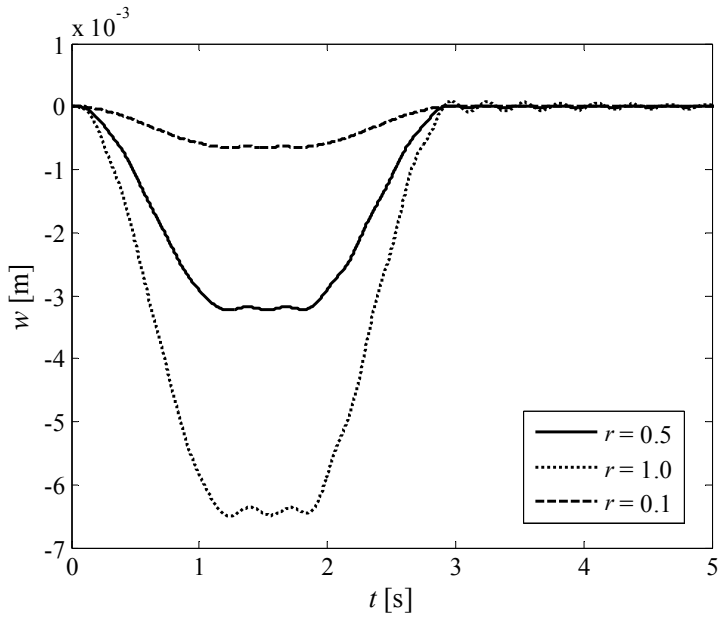


Figure 5. Effect of parameter r on $w(t)$.

Effect of the length of the load

Parameter values are assumed as in the reference case with $v = 40 \text{ m/s} = 144 \text{ Km/h}$, varying μ from 0.1 to 2. Response functions $w(t)$ and $H(\alpha)$ are computed in $x = L_b/2$, $y = l_b/2$, as reported in Figs. (7) and (8) respectively. Parameter μ (related to L_t) is able to significantly affect the behaviour of $H(\alpha)$. Note that the plots of $H(\alpha)$ in the case $\mu \geq 1$ are superimposed.

Effect of structure damping

Parameter values are assumed as in the reference case, varying ζ from 0 to 1. Response function $H(\alpha)$ is computed in $x = L_b/2$, $y = l_b/2$, as reported in Fig. (9). Raising ζ reduces the amplitude of oscillation of $H(\alpha)$, until its behaviour becomes monotonic (however this is not the case for real bridge structures). On the contrary, varying the frequency parameter β within the range of real bridges scarcely affects the behaviour of $H(\alpha)$, since the plots are almost superimposed [8].

Effect of partially distributed load

Parameter values are assumed as in the reference case, with $L_t = 24 \text{ m}$ and $v = 40 \text{ m/s} = 144 \text{ Km/h}$. Different loading distributions are compared: the continuous one, as represented in Fig. (1), and ‘partial’ distributions consisting of two shorter sections in which the load is distributed. The assumed ‘partial’ distributions are given by:

$$0 \leq \xi \leq \frac{L_t}{\chi} \quad \text{and} \quad \left(L_t - \frac{L_t}{\chi} \right) \leq \xi \leq L_t \quad (16)$$

with $\chi < 0.5$ ($\chi = 0.5$ yields the ‘continuous’ distribution). Since for the continuously distributed load it is assumed $r_0 = 0.5$, in the case of the ‘partially distributed’ load described by Eqn. (16) r_0 increases to $r = 1/(2\chi) \times r_0$. Response functions $w(t)$ and $H(\alpha)$ are computed in $x = L_b/2$, $y = l_b/2$ for different values of χ (1/6, 1/48, 1/480) as reported in Figs. (10) and (11). Load distribution variations such as that described in Eqn. (16) may dramatically affect the behaviour of the response function $H(\alpha)$.

Effect of time dependent matrices

Parameter values are assumed as in the reference case, with $v = 40 \text{ m/s} = 144 \text{ Km/h}$. The effect of neglecting the time dependent matrices $\Delta \mathbf{M}$, $\Delta \mathbf{C}$, $\Delta \mathbf{K}$ on the solution $w(t)$ is evaluated by introducing a relative error, according to:

$$\varepsilon_M = \frac{w(t) - [w(t)]_{\Delta \mathbf{M}=\mathbf{0}}}{\max_t |w(t)|} \quad (17)$$

where $[w(t)]_{\Delta \mathbf{M}=\mathbf{0}}$ refers to the solution computed assuming $\Delta \mathbf{M} = \mathbf{0}$ in Eqn. (13). Similarly, ε_C and ε_K can be defined, considering $\Delta \mathbf{C} = \mathbf{0}$ and $\Delta \mathbf{K} = \mathbf{0}$. The error functions ε_M , ε_C and ε_K are plotted versus time in Fig. (12), where ε_{Tot} represents the total error, assuming $\Delta \mathbf{M}$, $\Delta \mathbf{C}$ and $\Delta \mathbf{K}$ equal to $\mathbf{0}$ at the same time. The smallest, and negligible contribution to the error ε_{Tot} appears to be ε_C , while the main contribution is due to ε_M .

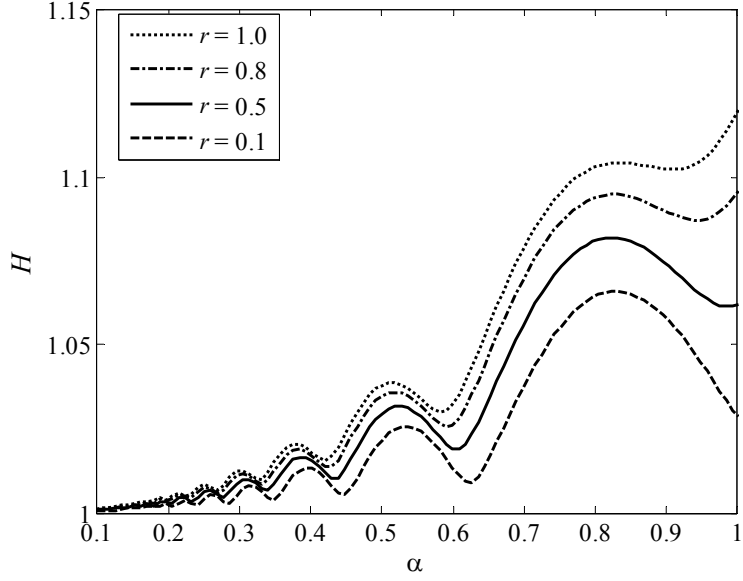


Figure 6. Effect of parameter r on $H(\alpha)$.

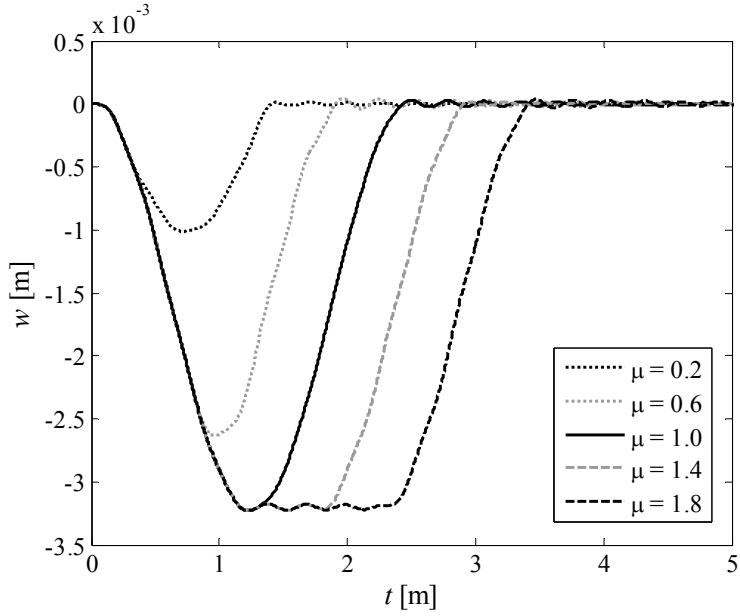


Figure 7. Effect of parameter μ on $w(t)$.

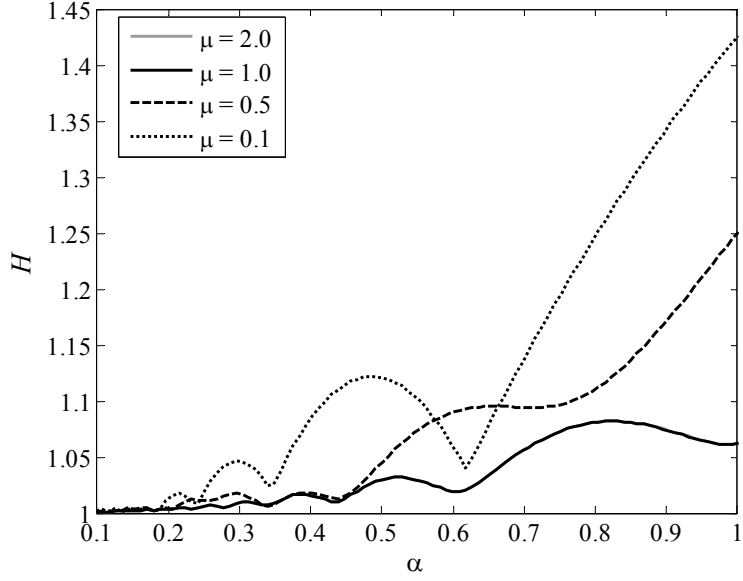


Figure 8. Effect of parameter μ on $H(\alpha)$.

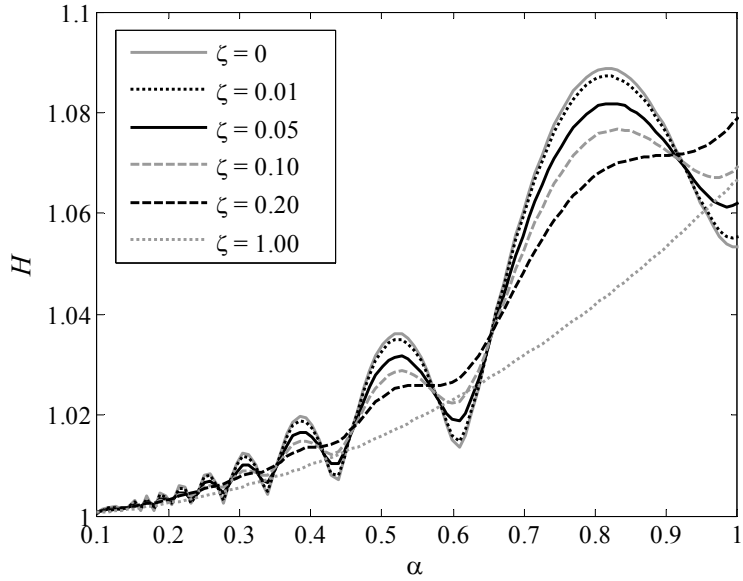


Figure 9. Effect of parameter ζ on $H(\alpha)$.

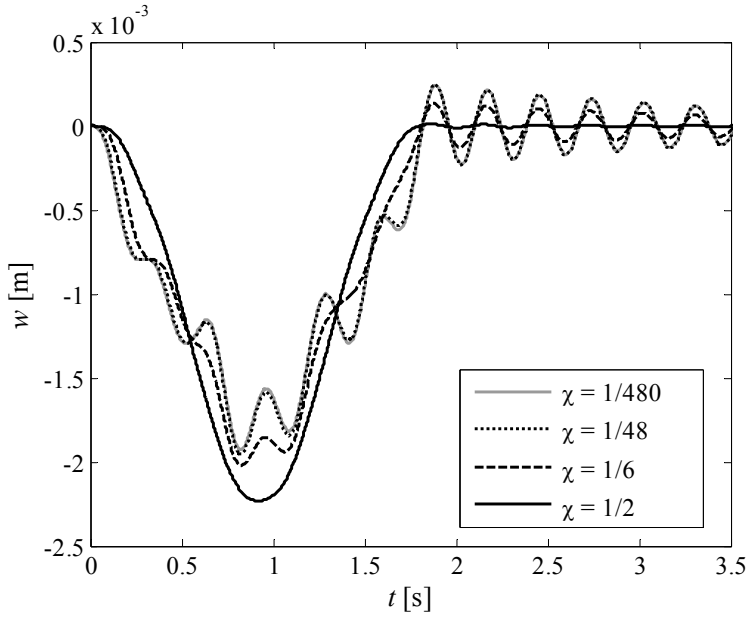


Figure 10. Effect of partially distributed load on $w(t)$.

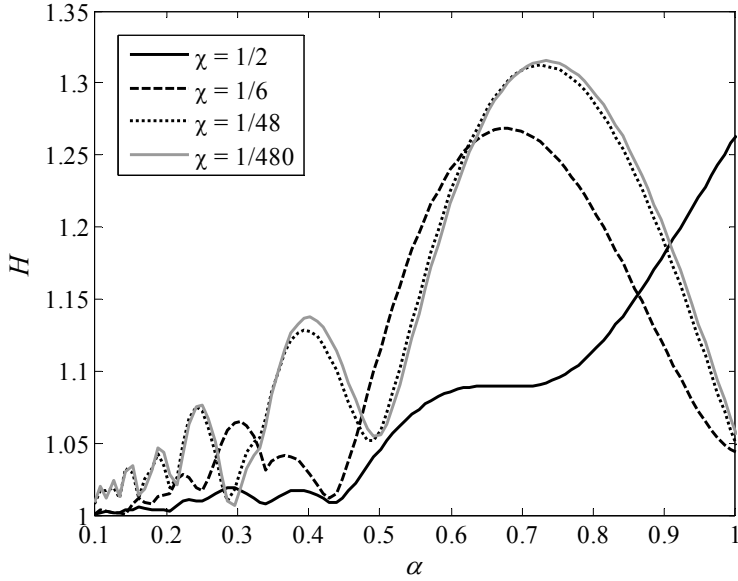


Figure 11. Effect of partially distributed load on $H(\alpha)$.

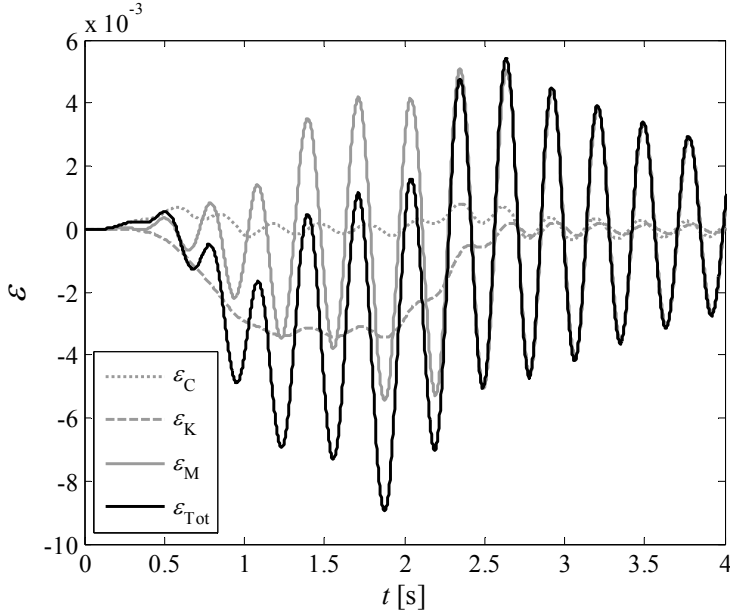


Figure 12. Effect of neglecting time dependent matrices $\Delta\mathbf{M}$, $\Delta\mathbf{C}$, $\Delta\mathbf{K}$ on $w(t)$.

4. DISCUSSION

Function $H(x, y; \alpha)$ appears to be an effective tool for studying the dynamic behaviour of a structure crossed by travelling loads with constant speed, in some way equivalent to a ‘frequency response function’ for time-varying coefficient systems. This function shows peculiar undulating trends, influenced by the parameters governing Eqn. (13). The response can be evaluated at any coordinate point (x, y) of the plate, making it possible to study the variation of structural deflection also along the y coordinate.

Mass parameter r can produce important shifts in magnitude, but not in shape, as shown in Figs. (5) and (6). On the contrary, length parameter μ controls both shape and magnitude of $H(\alpha)$, but only in the case $0 < \mu \leq 1$, as shown in Figs. (7) and (8).

The damping parameter ζ has the effect of progressively smoothing the oscillation of $H(\alpha)$, until it becomes monotonic (though the latter limit case is not realistic for actual bridge structures): in general, the reduction in amplitude becomes particularly significant at high speed, as shown in Fig. (9). Frequency parameter β , within the range of real bridges, scarcely affects the behaviour of $H(\alpha)$, so H may be considered independent from β [8].

Changes in the spatial distribution of the load can produce dramatic variations in $H(\alpha)$, shown in Figs. (10) and (11). This behaviour should highlight the importance of properly modelling the ballast, directly influencing the load distribution on the actual structure.

The contribution to the solution of the time dependent matrices $\Delta\mathbf{M}$, $\Delta\mathbf{C}$ and $\Delta\mathbf{K}$ is globally not negligible, however the effect of $\Delta\mathbf{C}$ is usually very small in comparison with the contributions of $\Delta\mathbf{K}$, and especially of $\Delta\mathbf{M}$.

5. CONCLUSIONS

In this study, the dynamical behaviour of railway bridges crossed by travelling trains was investigated by adopting a simplified model, i.e. a plate loaded by a travelling distributed mass, solved by means of the Rayleigh-Ritz method.

The effects of each of the model governing parameters was studied introducing a dynamic function of the travelling speed, equivalent to a 'frequency response function' for time-varying coefficient systems. This function can be an effective tool for studying the dynamic behaviour of a structure crossed by travelling loads, since the travelling speed is the most important parameter influencing the dynamic stresses in railway bridges, which in general increase with increasing speed.

In particular, it was shown how different spatial distributions of the load can deeply influence the dynamic response of the structure, highlighting the importance of properly modelling the ballast. Future work will thus concern this significant problem.

Acknowledgements

This study was developed within the INTERMECH laboratory with the contribution of the Regione Emilia Romagna - Assessorato Attività Produttive, Sviluppo Economico, Piano telematico, PRRIIT misura 3.4 azione A Obiettivo 2.

REFERENCES

- [1] Fryba, L., 1999. *Vibration of Solids and Structures under Moving Loads*, 3rd edition, Telford.
- [2] Fryba, L., 1996. *Dynamics of Railway bridges*, Telford.
- [3] Lin, Y.H., 1994. "Vibration analysis of Timoshenko beams traversed by moving loads". *Journal of Marine Science and Technology*, **2** (4), pp. 25-35.
- [4] Stancioiu, D., Ouyang, H., Mottershead, J.E., 1999. "Vibration of a continuous beam excited by a moving mass and experimental validation". *Journal of Physics*, Conference series **181**.
- [5] Adetunde, I.A., 2007. "Dynamical Behavior of Euler-Bernoulli Beam Traversed by Uniform Partially Distributed Moving Masses". *Research Journal of Applied Sciences* **2** (4), pp. 476-483.
- [6] Lin, Y.H., Cho, C.H., 1993. "Vibration suppression of beam structures traversed by multiple moving loads using a damped absorber". *Journal of Marine Science and Technology*, **1** (1), pp. 39-48.
- [7] Timoshenko, S., Young, D.H., Weaver, W., 1974. *Vibration problems in engineering*, 4th edition, Wiley.
- [8] Catania, G., Sorrentino, S., 2010. "Dynamic analysis of railway bridges by means of the spectral method". In *Proceedings of Vibrations, Chocs & Bruits XVII*, 2010 Lyon, France.

FULL-FIELD OPTICAL MEASUREMENTS FOR ADVANCED STRUCTURAL DYNAMICS: FIRST OUTCOMES

Alessandro Zanmarini

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: a.zanmarini@unibo.it

Manfred Neumann

Institute of Mechanics and Mechatronics,

Vienna University of Technology, Austria

E-mail: manfred.neumann@tuwien.ac.at

Johann Wassermann

Institute of Mechanics and Mechatronics,

Vienna University of Technology, Austria

E-mail: johann.wassermann@tuwien.ac.at

Abstract. *A growing activity of the first author on full-field vibration measurements and promising analyses on dynamic characterisation of components, on dynamic strain & stresses and on cumulative damage maps has drawn the attention of a foreign institution such as Vienna University of Technology up to offer full access to its Schwingungs- und Strukturanalyse / Optical Vibration Measurement Laboratory, equipped with state-of-the-art instruments implementing different & complementary approaches, in order to give the researcher the unique opportunity to extend the stream of researches in the strategic field of image-based measurement technologies. The early activity carried on in the frame of the settled Full-Field Optical Measurements for Advanced Structural Dynamics project is discussed in detail, showing the first achievements.*

Keywords: *full-field optical measurements, SLDV, ESPI, DIC, vibration*

1. INTRODUCTION

The structural dynamics community has seen the advent of optical full-field techniques since the time of silver halide holography that, with pulsed or stroboscopic laser light, permitted the acquisition of complex displacement patterns in dynamic events. Holography required the development of chemical paper or recording compound, meaning that a whole dynamic analysis in a wide frequency range was not competitive with growing piezoelectric transducers like accelerometers [1]. But at the same time, it gave a valuable insight of the shapes of the surfaces in the spatial-frequency domain, then of basic importance in determining the

best location of lumped sensors in challenging experiments. Later the spectrum analysers allowed multi-channel acquisitions [2], extending the number of transducers up to hundreds and coming back to sense spatial patterns by means of grids of lumped sensors, even if quite invasive when dealing with light structures. When the Laser Doppler Vibrometer reached scanning functionality (SLDV), this grid of sensing locations became fine, non-invasive and fully compatible with all the sampling and analysis techniques developed up to that point, shortly becoming a benchmark in structural dynamics for its ability of acquiring quantitative vibration patterns in a wide frequency range. But the enhancements in electronics and integrated circuits that permitted these milestones in measurements techniques were also implemented in computers, fast memories and CCD sensors, a grid of light sensitive devices (photosites) that could sample synchronously the field of view of an optical lens. The silver halide layer was then substituted by the new light recording media, of inferior quality, but ready digital and with a fast processing. The Electronic Speckle Pattern Interferometry (ESPI), also known as Digital Holography [3], moved its first steps allowing displacement patterns also of moving objects, when frozen between two instants of their motion by means of a pulsed or stroboscopic laser light, even if the first CCD did not allow high speed recordings, because with that lighting system any motion appeared as static deflection to the camera reference. Then it was the time of stereoscopic full-field techniques, working on correlated digital images coming from two cameras pointing to the same field of view in white light, leading to a three-dimensional sensing of the surface, but also of its behaviour in successive acquisitions. While Digital Image Correlation (DIC) started to be applied on static deformation problems, it is full compatible with time histories of dynamic events, just limited by the performances of the electronics and calculation time in reconstructing the displacement fields. In this way, DIC can be seen as a spatially consistent and really simultaneous grid of sensors acquiring time-dependant signals.

Image-based inquiries are then gaining interest in the structural dynamics community with their growing potentialities coming from the recent advancements of camera and computing electronics. Cameras that are more sensible to light and allow greater recording speeds and buffers are more and more progressively hitting the market, raising expectations on the developments that may come in sensing dynamic motions of surfaces. Nevertheless, the equipment that applies optical techniques to mechanical measurements still lacks a consolidate background of practice & analysis procedures and comparative studies, especially in the area of dynamic measurements such as those for vibration patterns. Today the state-of-the-art of full-field technology can deliver displacement maps at relevant frequencies for dynamic ESPI or in time-histories at low-medium frame rates for Hi-Speed DIC, while SLDV reaches the target of addressing FRFs quantities with high spatial resolution. But from the high consistency of full-field displacement patterns can come relevant quantities like dynamic strain & stress maps, together with refined eigenvectors, once extracted from the huge amount of the full-field datasets. Full-field techniques, in spite of their potentialities, do not find an adequate implementation in noise & vibration studies or in challenging designs.

The project *Full-Field Optical Measurements for Advanced Structural Dynamics* aims at closing the gap between the theoretical potentialities of full-field instruments and the actual quantities they can measure, with particular attention towards NVH and structural dynamics requirements. This project has its strength in a comparative approach between SLDV, ESPI and DIC. After having introduced the path that has brought to this challenge, the activity done at TU-Wien will be described with details and commented for the future developments.

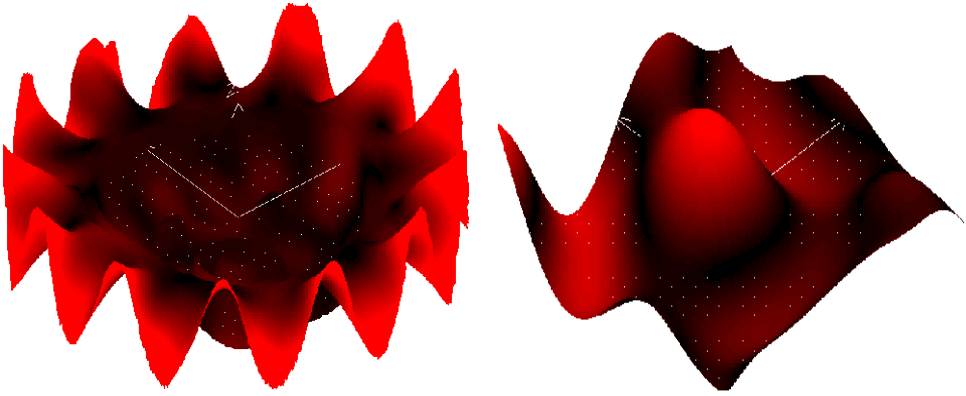


Figure 1: Acquisition examples of complex-valued operative deflection shapes of a vibrating brake disk at 16120Hz and of a plate at 948Hz

2. PREMISES

Past research activities on full-field measurements

Due to lack of funds in supporting his research at Bologna University during 2004, Dr. Zanarini was forced to search for other research opportunities abroad. There came to his attention the vacant post doctoral position in the European Commission - Marie Curie Industry Host Fellowship *Speckle Interferometry for Industrial Needs* project, held at Dantec Ettmeyer GmbH, Ulm (Germany), aiming at enhancing the vibration measurements in a full-field domain, by means of ESPI technologies. The past experiences on spatial sampling errors in structural testing during his European Doctorate on Sound and Vibration Studies at Katholieke Universiteit Leuven, Belgium, gave Dr. Zanarini the challenge that his work in the framework of the *Speckle Interferometry for Industrial Needs* project, by means of extensive measurements and analysis software, would bring to promising issues. After successfully obtaining the fellowship, from April 2004 to July 2005 there was the possibility for Dr. Zanarini to broaden his skills and competences on full-field technologies, like pulsed and stroboscopic ESPI and Shearography, and on analysis software, which joined the traditional NVH approaches with the high accuracy in spatial domain of the full-field measurements; these vibration measurements were focused on a damaged composite panel [4] and on a brake disk dynamic characterisation [5].

After the Marie Curie fellowship period, once returned to Bologna University, Dr. Zanarini developed the analyses of the acquired full-field vibration measurements, spreading the results thereof on complex-valued displacement fields in frequency domain as shown in Fig.(1), full-field frequency response functions and structural dynamics [6, 7], surface dynamic strain & stress, cumulative damage maps and fatigue life assessment [8, 9, 10] as in Fig.(2).

The invitation by TU-Wien and the project

In September 2008 Prof. Wassermann, from the Institute of Mechanics and Mechatronics, Machine Dynamics and Measurement Technologies, Vienna University of Technology, Aus-

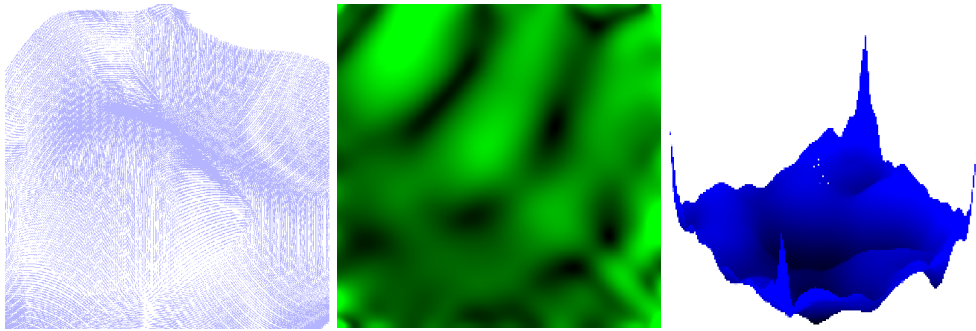


Figure 2: Evaluated quantities on the vibrating plate: first principal dynamic strain map at 948Hz, Von Mises equivalent dynamic stress map at 948Hz, cumulative damage distribution for white noise excitation spectrum [9]

tria, invited Dr. Zanmarini to spend a research period abroad at the optical measurement laboratory of TU-Wien, thanks to the promising results shown in the presentation *ESPI measurements in structural dynamics: fatigue life assessment* [10]. The invitation consisted in offering full access to the well-equipped full-field vibration measurement lab in order to let Dr. Zanmarini freely widen his research streams and analysis approaches.

In December 2008 Dr. Zanmarini obtained financial support from DIEM, Mechanical Engineering Dept., University of Bologna - Marco Polo Program, for the first four months of the living expenses abroad, which enabled him to start the project in partnership with TU-Wien.

On January 2009 Dr. Zanmarini started the period as visiting professor (Gastprofessor) at TU-Wien on the project *Full-Field Optical Measurements for Advanced Structural Dynamics*, with the following research topics:

- comparison of different full-field technologies;
- complex valued full-field measurements, modal superposition, role of damping;
- dynamic strains and stresses, fatigue tests and predictions by means of stress and strain based approaches;
- damage detection by means of displacement field processing;
- full-field FRF for dynamic analysis, NVH and EMA;
- full-field model updating, vibro-acoustic modelling;
- feasibility studies on Experimental CMS and Full-Field Modal Analysis.

3. THE LABORATORY AT TU-WIEN

Vienna University of Technology hosts a well-equipped laboratory for optical vibration measurements, with instrumentations at the state-of-the-art. The peculiarity of this lab resides not only in the measurement equipment, later reported, but also in the infrastructure itself. The measurement environment is located in a dedicated room with seismic floor to filter out the

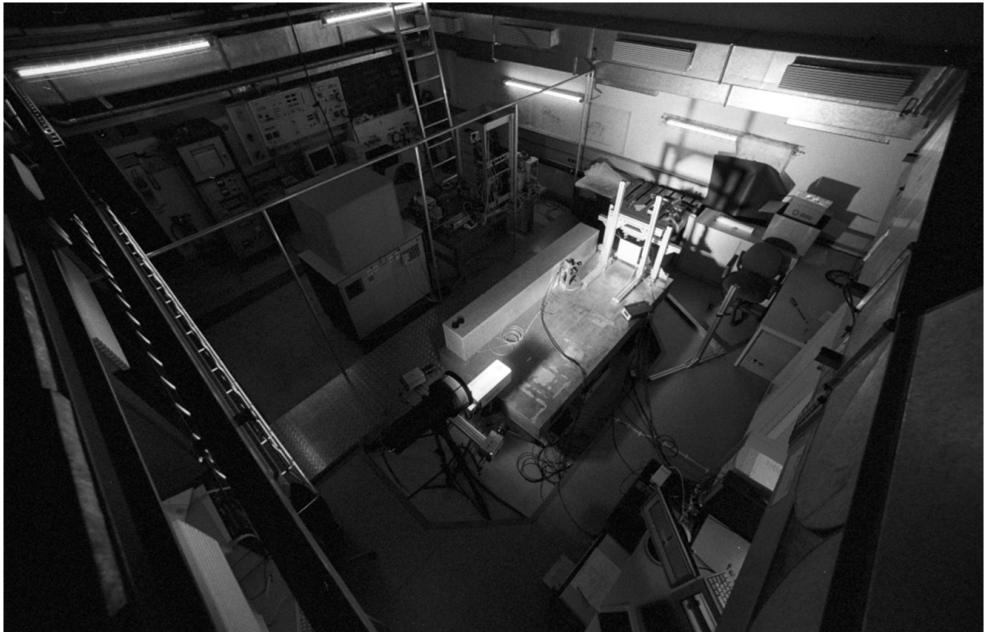


Figure 3: The optical vibration measurement laboratory at TU-Wien

building and ground influences, as shown in Fig.(3); plus, the set-up can be arranged on an air-spring anti-vibration optical table, to assure an inertial reference. Moreover, even if the laboratory is placed in the basement of the TU-Wien building, it is reachable in few seconds from the offices, which is a relevant matter in long and continuous measurement campaigns, meaning that facilities' layout is of uttermost importance in competitive research. Mechanical and electronic workshops with competent technicians are at total disposal to complete the services offered to the researchers, and they both contributed to the success of the project, because machined parts were needed for the set-up and auxiliary electronics was made for the triggering of the acquisition systems.

Traditional tools for vibration & modal analysis can be found at TU-Wien, like acquisition systems (an LMS-Scadas Mobile SCM05 16 channels acquisition unit was used, with its dedicated LMS Test.Lab software suite), power amplifiers, function generators, shakers, hammers, impedance heads, software, etc.. All these traditional tools were exploited at their best to set-up the experiments and to tune them for the best results.

The optical vibration measurement equipment available at TU-Wien consisted in:

- 1D Scanner Laser Doppler Vibrometer (SLDV), Polytec PSV 300, with 1D (out-of-plane) scanning head OFV-056;
- 3D Electronic Speckle Pattern Interferometry (ESPI), Dantec Ettemeyer Q-500 VibroESPI with quantitative amplitude & phase analysis, high resolution camera sensor (1380 x 1035 pixels) and stroboscopic laser light system (5Watt cw - 532nm);
- 3D Hi-Speed Digital Image Correlation (DIC), Dantecdynamics Q-450 and hi-speed

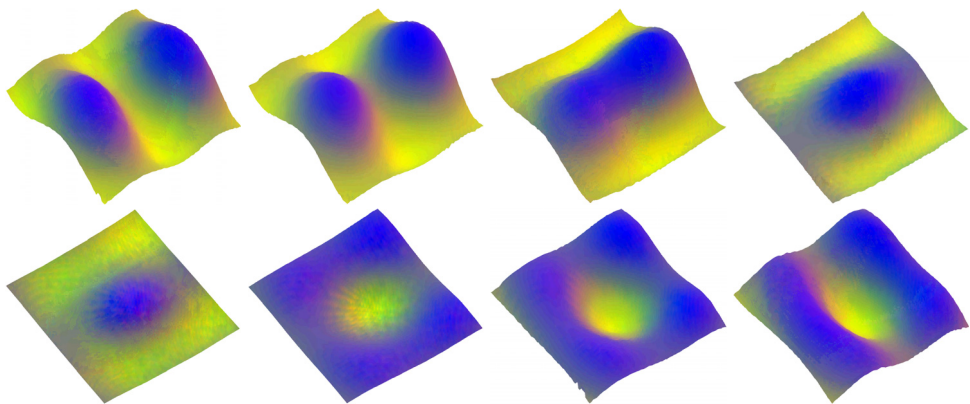


Figure 4: A time-step sequence of membrane vibrations, animated by the prepared software

cameras (Nanosense Mk III, 4 GB of memory - 1040 fps at 1280 x 1024 pixels maximal resolution) with hi-power & hi-frequency white light system.

All these facilities made it possible to set-up a unique comparison of the 3 different optical technologies in full-field vibration measurements, and to design experiments that can support further research in modal analysis.

4. THE ACTIVITY AT TU-WIEN

Learning new full-field measurement technologies

The past activity of Dr. Zanmarini had provided relevant skills in dealing with the ESPI systems, especially in stroboscopic light, for the sine stepped dynamic measurements. The set-up at TU-Wien was a chance to extend the measurements database with even more accurate measurements, by means of reduced frequency gap between acquisitions and more samples for the complex-valued representation. Extensive practice was made on the set-up to tune the restrains for the low frequency rigid body motions and the excitation source, in order to let ESPI system acquire data of enhanced quality with regards to those of the past.

On SLDV and DIC technologies Dr. Zanmarini had no previous experience, they being on the contrary new measurements technologies for him, and required extended & time-consuming practice to gain the necessary competence to proficiently exercise them to their best and to find the conditions under which to make the best compromise to run comparative tests between the different technologies possible. From a theoretical point of view, activity was spent to understand the principles and limits of each approach, and the needed acquisition parameter optimisation. From a lab-based point of view, in particular, relevant attention was put on the requirements of the DIC approach, in the terms of the random pattern to be applied by a spray paint coating on the object under test and of the calibration of the stereo camera system, to calculate all the intrinsic and extrinsic parameters of the lenses and perspective & point of view. The paint coating gave a rougher surface, which was perfect for ESPI by means of scattering light and speckles uniformly, and a compromise solution for SLDV, which in some location showed a lightly weaker signal quality, compared to the excellent signal acquired in most of the other parts of the surface. Relevant work was also done for

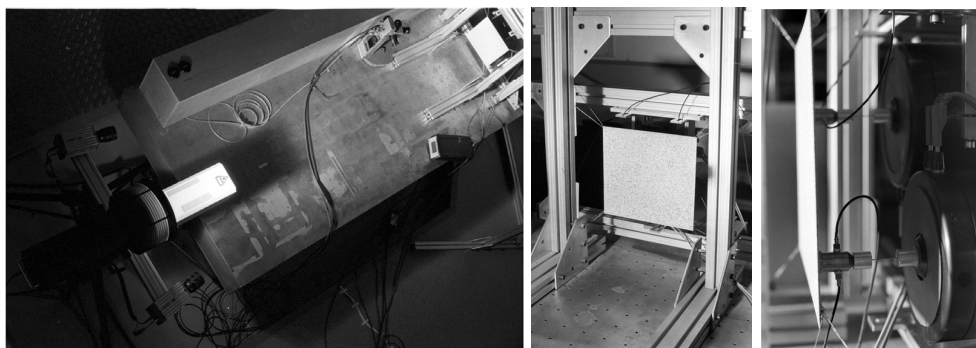


Figure 5: The common experimental set-up for vibration measurements by means of optical equipment

the acquisition of random events, by means of automatically triggering signals from the set-up. On SLDV, different trials were made to understand the limits on the spatial accuracy as compromise with measurement time and repeatability of the dynamic behaviour of the sample, when this is sensible to the thermal conditions of the environment. In fact, along the day non-uniform thermal heating affected the room and this was probably the cause of unexpected results on SLDV.

For the forecast future analysis, attention was paid to the possibility to read and elaborate the experimental results, by means of acquiring the knowledge in order to interface oneself with the dedicated software of the instruments and their proprietary file formats. Software was thus developed for data reading & representation, as can be seen in Fig.(4) where a sequence of hi-speed DIC displacement patterns is reported from the graphical interface developed. This was the early version of the tools that will be developed for further enhancements in this stream of researches.

Preparing a common experimental set-up

In a growing field, such as that of full-field optical measurement technologies, comparative studies become of uttermost importance in assessing the achieved quality and the consistency with previous approaches. Full-field optical instruments still need to demonstrate their potentialities and benchmarks are missing in dynamic measurements. It was quite a unique chance to have in the group of the available instruments at TU-Wien the SLDV, acknowledged as a reliable approach in structural dynamics [11] to spatially refine the vibration measurements of optically accessible surfaces. It was thus decided to set-up the experiments with the SLDV as common reference. Since the range of measurements can vary from one technology to the other, a linearity check was successfully performed on the test set-up that was prepared as common basis for the comparative studies. In fact, ESPI is the most sensitive of the three technologies, but limited in displacement range, and required minimal power in the input, while DIC is less sensitive and generally requires bigger displacements and hi-power in the input source; SLDV works in an intermediate range of displacements and input power.

Another requirement for the set-up was that of showing a high modal density and closeness of the eigenmodes in a frequency range common to all three optical techniques, to stress

the multi-input-*many*-output opportunity for extended analysis approaches [12, 11] and to investigate the independence of the vibration patterns from the specific excitation point. The conceived experiments can also be of relevant aid for further developments towards full-field modal analysis, extensive considerations on FRFs, dynamic strain-stress tensor modelling, and cumulative damage & fatigue life assessment.

For these reasons, a thin aluminium plate with the shape of a nearly perfect square was adopted. It was restrained by thin wires glued on the back at the four corners and fixed to pillars on the anti-vibration optical table, as shown in Fig.(5). The system of restrains had only to suppress the low-frequency rigid body motions and leave the plate in a quasi-free-free condition, to let appreciate the behaviour of the borders in dynamic motion. Two shakers were attached to the back of the plate and driven by the acquisition systems.

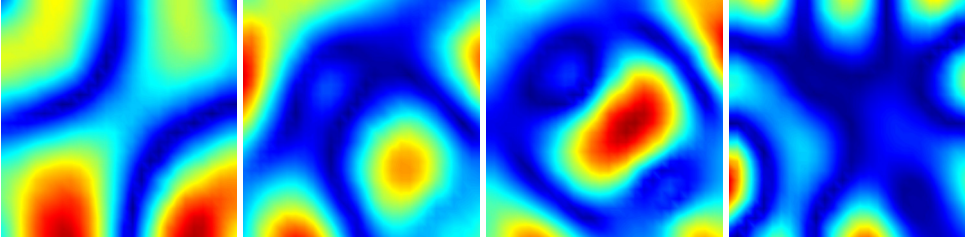
The tuning of the common experimental set-up was critical, especially in finding the right tension for the restraining wires, which were excited by so discrepant motions, from the sub-micron range while driving ESPI measurements to the sub-millimetre range in DIC. In spite of difficulties, the right adjustment was reached, to assure the wanted linear behaviour and avoid undesirable disturbing vibrations of the wires. The acquired data can be eventually compared also with those from ESPI measurements collected in 2005 [6].

All three measurement systems were placed in front of the plate, with acquisition direction orthogonal to the flat surface and the centre of the plate as reference point for the alignment. The field of view was maximized in each system to retain the maximal spatial resolution available, according to the dynamic requirements. Due to these latter, the DIC pixel resolution was reduced to let the hi-speed cameras run at higher frame-per-second rates.

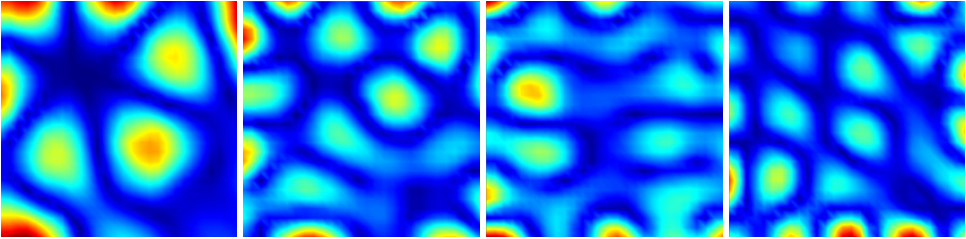
5. FIRST OUTCOMES

During this early part of the project *Full-Field Optical Measurements for Advanced Structural Dynamics*, an intensive lab-based activity was carried out to extend the skills on full-field vibration measurements, regarding refined dynamic ESPI data acquisition, SLDV and Hi-Speed DIC new competences. The chance to have at TU-Wien different and complementary technologies, like ESPI and SLDV in frequency/space- and hi-speed DIC in time/space-domains, gave the unique possibility to further develop sensing-approach-independent ideas on full-field measurements, disregarding the used technique. Consequently, the early version of new numerical tools was written for a deep data analysis, awaited in the near future.

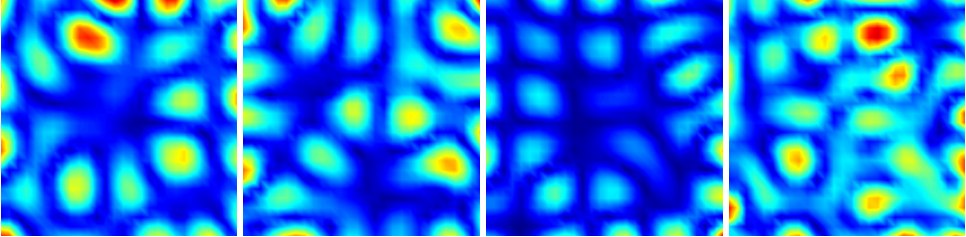
But in order to prepare the measurements for the next challenges, it was necessary to conceive and manage innovative & complex experiments, with a relevant work in the lab and on the computing facilities of TU-Wien, mixing the knowledge in experimental analysis [12, 11] with the practice on the available optical techniques, pushed to their performance limits in terms of accuracy and repeatability of results. The conceived common set-up shown in Fig.(5) permitted the full-field vibration measurements, in a multi-input-*many*-output general scheme. The acquisitions were assured on the linearly behaving plate, in the same overlapping frequency range of 20-1000Hz for all three techniques, with high frequency resolution, below 1Hz where possible. ESPI lower frequency limit is set to 20Hz in the acquisition system electronics; the hi-speed cameras used for DIC could not go over 2000 fps, which was the sampling frequency in this image-based case. Further, for ESPI and SLDV the multi-input-*many*-output comparison was extended to higher frequencies up to 3kHz. At testing time the spatial resolution on the plate side was not made uniform for the three technologies, growing from SLDV (up to 60 dofs) to DIC (about 120 dofs) to ESPI (about 500 dofs).



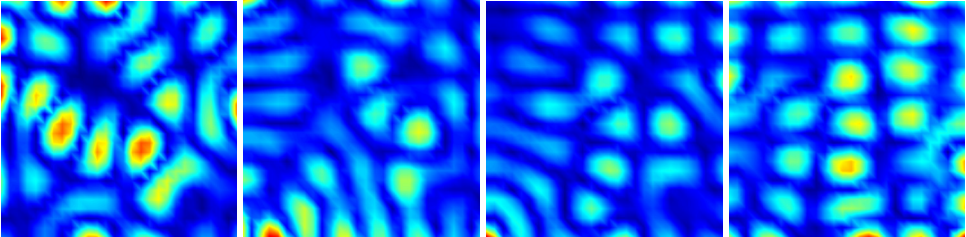
a) measurements at: 214Hz, 323Hz, 437Hz, 633Hz



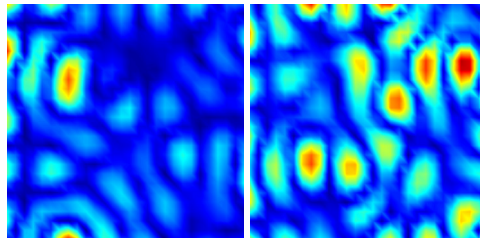
b) measurements at: 708Hz, 1276Hz, 1425Hz, 1698Hz



c) measurements at: 1852Hz, 1948Hz, 2228Hz, 2433Hz



d) measurements at: 2663Hz, 2794Hz, 2841Hz, 2900Hz



e) measurements at: 3056Hz, 3112Hz

Figure 6: SLDV mobility measurements from input 1

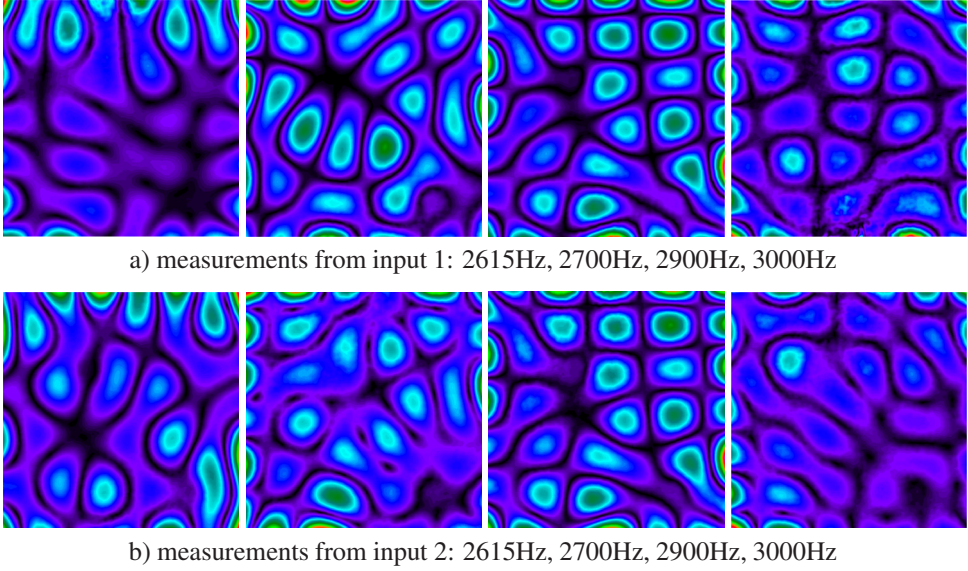


Figure 7: ESPI displacement measurements from two different inputs

SLDV gives the wanted measurements like well-known accelerometers, except that it measures surface velocity instead of acceleration, having the same requirements of time-per-location as a single lumped sensor, with regards to the desired frequency accuracy, thus requiring a lot of time in scanning the surface with high resolution (as quoted above, up to over 3000 scanned locations). Some mobility amplitude maps are shown in Fig.(6) for SLDV at some frequency lines, with relatively low spatial accuracy. SLDV is not a native full-field technology, which acquires instead all the data together from each location: then it requires that the object under test behaves in steady-state conditions. But, since a full acquisition with high spatial resolution lasted also a few days, the cited room heating problems were faced in the tests, meaning that the non-synchronous measurements showed unexpected inconsistency on the spatial domain, like lost phase correlation or time delay occurred at specific time-steps, leading to mistaken bands / behaviours on the measurements, which needed to be rescanned. Other causes could be overheating and clock errors in the electronics of the whole system, due to the continuous work, but it was not possible to find a rapid solution.

Each frequency step in ESPI measurements needed about 2 minutes to acquire and evaluate the complex-valued displacement maps; this was repeated for all the frequency lines of interest. To assure the best quality of the resolvable fringe systems, with such a sensitive system as ESPI that works in the sub-micron range, a very low input force was calibrated & optimised at each narrow frequency interval of interest, in order to obtain the best signal-to-noise ratio and avoid unreadable areas in the fringe maps, coming from exceeding displacements where the fringes are no more clear and distinguishable in the interference patterns. Some deflection shapes can be seen in Fig.(7), showing the high quality displacement amplitude fields induced by the two different input locations, which deserve further investigations.

The acquisition of dynamic signals was very fast by DIC, like acquiring one measurement

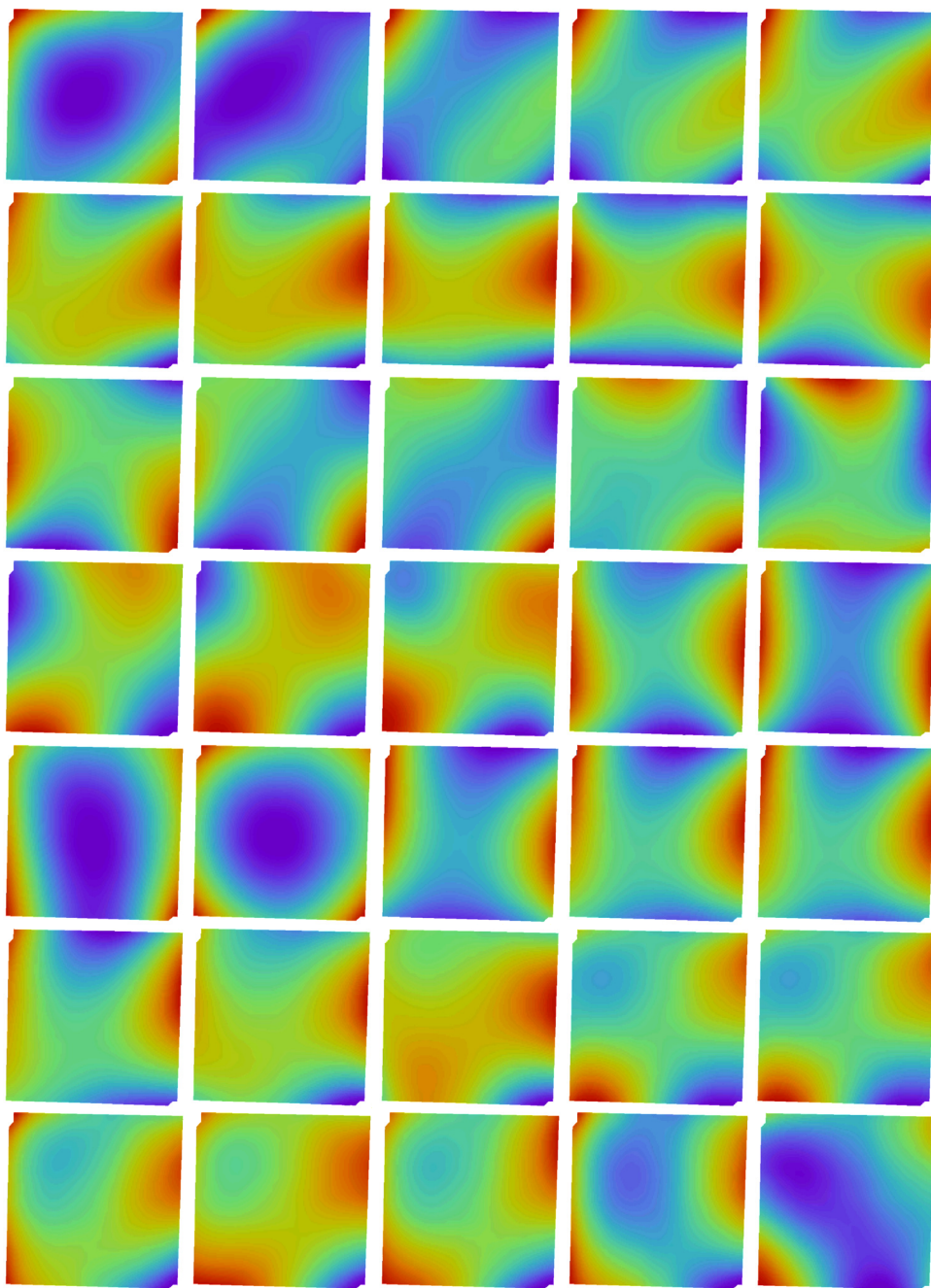


Figure 8: DIC displacement measurements from input 1 at sequential frames in time domain, part I

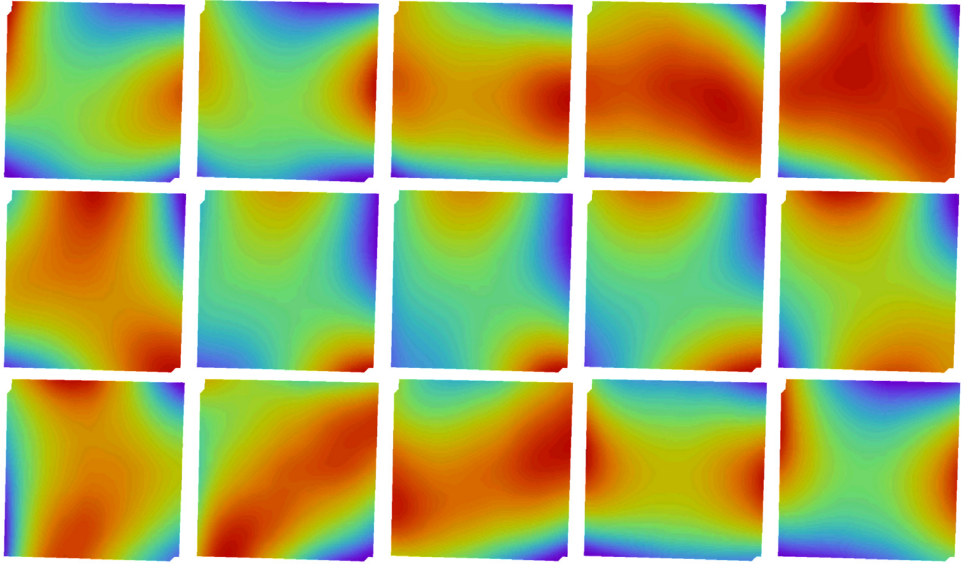


Figure 9: DIC displacement measurements from input 1 at sequential frames in time domain, part II

from a single lumped sensor such as a common accelerometer, but the off-line work was quite heavy: half an hour was required to download the DIC time series from the cameras, about 3 days to evaluate the corresponding displacement fields in the time domain. Strong power was needed by the shakers, due to the weaker sensibility to displacement of the DIC technology, compared to ESPI and SLDV. In fact, at TU-Wien DIC was intended for much larger displacements in the range of millimetres, but proved in this project to be able, once precisely calibrated, to sense quite small vibrations also in the sub-millimetre range. The quality obtained from the measurements is promising for next elaborations. In Fig.(8,9) are shown in row order some full-field deflection shapes from DIC at different & increasing time steps, with the time-dependent modal superposition result of the lower eigenmodes in evidence, highlighting the continuous mixing of the modal base vectors to generate complex displacement patterns with promising spatial consistency.

The structural dynamics retained in the measurements seems again to play a relevant role in determining the complex shaped distribution of flexural motion inside the surface area and reflect that high modal density searched for at the time of the design of the experiment.

All the measurement systems available worked in a single-input-*many*-output scheme, since it was not possible to drive the two shakers together and evaluate the transfer functions from a unique multi-input-*many*-output test, even if the LMS Test.Lab software suite & Scadas front-end could have generated the right inputs to both shakers. Thus, to overcome this limitation and have a real multi-input-*many*-output dataset, separate tests were run for each shaker input.

There resulted a wide experimental dataset useful to investigate the influence of acquisition parameters on the three different techniques and on the measurable quantities, based on

the comparison of different optical technologies on the same inquiry domain, also with the chance to draw statistical considerations.

A broad band frequency analysis of the dynamic behaviour of the thin plate with full-field measurements and high spatial resolution will shortly come by means of receptance maps, to assess the potentialities of these spatially accurate measurements in extending the modelling consistency with real complex test rigs. The receptance maps permit the experimental modelling of the component as subjected to a unitary amplitude input force in a wide frequency domain. From high quality receptance maps can come the evaluation of the dynamic strain & stress tensor models and fatigue life estimations: these results will prove the advantages of having such complex instrumentation as that needed for full-field dynamic measurements.

An increased experience about algorithms for full-field vibration analysis also followed, which can address the full-field datasets as sensing-system independent and further investigate the physics of the underneath mechanical system with more reliability.

The technical exchange with *Dantecdynamics GmbH*, to refine their software in processing wide DIC datasets, extended the contacts with potential partners in applications of this stream of research.

6. TEACHING ABROAD

The researches held at TU-Wien also gave the opportunity to attract Master and Doctorate students to the topics that stay behind the project *Full-field optical measurements for advanced structural dynamics*. Two optional courses in English were offered to determined classes of 5-7 students, also from other institutes and disciplines, and coordinated with the Institute of Mechanics and Mechatronics: the first about the traditional structural dynamics advanced approaches, like Component Mode Synthesis, the direct/indirect characterisation by means of the impedance coupling method (FRF-based Substructuring & Transfer Path Analysis), named *Complementary approaches in structural dynamics and experimental modelling*; the second about system identification techniques and traditional modal analysis approaches & algorithms, extended to vibro-acoustic domain, named *Advanced techniques in applied modal analysis and vibro-acoustic modelling*. They were a good chance to share with the audience the many aspects to be known in the theoretical background of the current researches, in order to be able to understand the decisions that brought to the specific set-up. The courses permitted also the transfer of knowledge towards the personnel of the division of Machine Dynamics and Measurement Techniques of the Institute of Mechanics and Mechatronics at TU-Wien.

Future cooperative educational projects may come from this previous teaching activity, in the areas of Master and Doctorate studies, but also as seminar-based actions, with the aim of spreading the interest in advanced full-field experimental techniques in structural dynamics and advanced design procedures: students in the course on modal analysis techniques showed interest in cooperative education. There might come projects like Erasmus ones, between the institutions of this paper's authors, or even with more partners involved. This teaching activity can be the occasion to tight together not only academia, but also industrial partners, in order to strengthen the full-field techniques community and let it have more chances to gain the attention and funds of the European Commission - Framework Programs. In this contest, the project *Towards Experimental Full-Field Modal Analysis* (acronym TEFFMA) has been submitted to obtain FP7-People-Marie Curie Action-Intra-European Fellowship funds.

7. CONCLUSIONS

The invitation that started this relevant interdisciplinary joined research project is an important acknowledgement from a foreign institution for the researches carried out by Dr. Zanarini in his past career. But most of all, it offered the chance to broaden the inquiries on this basic research stream towards new challenging targets that only an extensive, broad, unique and promising experimental enquiry in a growing area can assure. The teaching opportunities on the research topics constitute thereby other ingredients in the fruitful spread of these advanced design and analysis themes.

With respect to the past achievements [7], the project *Full-Field Optical Measurements for Advanced Structural Dynamics* has determined a consistent refinement of experimental and analysis techniques, some already faced (like dynamic ESPI and traditional vibration measurements), others of new practice (like SLDV and Hi-Speed DIC). Moreover, an intensive study of the algorithms of data & modal analysis was performed, with the aim of making a better experimental set-up, ready for further developments in the analysis stage. Promising qualitative comparison between different full-field technologies & complementary domains will be performed in the next stage of the project. While the first comparisons will be made on full-field FRF measurements, the conducted experiments will be of great value also in other advanced design procedures and in proposing refinements of well-established technologies and numerical approaches, also in suggesting the establishment of new quantities like dynamic strain&stress tensor models and spectral full-field cumulative damage schemes.

As further developments of this activity, quantitative comparisons of the acquired data fields are planned in the near future. The assessment of the potentialities of the different technologies for the traditional studies like NVH approaches, Full-Field FRFs, broad band vibration characterisation and complex design challenges is felt as of uttermost importance in increasing the interest towards these full-field technologies by structural analysts. During this work of quantitative comparison, advancements in the processing algorithms and approaches will be sought, together with international research opportunities (e.g. EC-FP7) and cooperative actions to spread these technological results in the structural dynamics community.

Acknowledgements

Vienna University of Technology, in the person of Prof. Wassermann and his staff, are greatly acknowledged for the opportunity given to Dr. Zanarini to have full access to the *Schwingungs- und Strukturanalyse* / Optical Vibration Measurement Laboratory, during the basic research project *Full-Field Optical Measurements for Advanced Structural Dynamics* and teaching courses held in Vienna in 2009. Bologna University and its Mechanical Engineering Department are acknowledged for the financial support of the early stage of this research abroad. *Dantecdynamics GmbH* (Ulm, Germany) is also acknowledged for the precious technical support given. This work is intended to be part of the spreading of the results thereof.

REFERENCES

- [1] Walter, P. L., 2007. The history of the accelerometer 1920s-1996 prologue and epilogue, 2006. Tech. rep., Sound and Vibration Magazine, www.SandV.com, January. 40th anniversary issue.
- [2] Deery, J., 2007. The real history of real-time spectrum analyzers a 50-year trip down memory lane. Tech. rep., Sound and Vibration Magazine, www.SandV.com, January. 40th anniversary issue.

- [3] Á. F. Doval, 2000. "A systematic approach to tv holography". *Measurement Science and Technology*, **11**(1), pp. R1–R36. IOP Pub Ltd.
- [4] Zandarini, A., 2005. "Damage location assessment in a composite panel by means of electronic speckle pattern interferometry measurements". In Proceedings of the IDETC/CIE ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, Long Beach, California, USA, September 24-28, ASME, pp. 1–8. Paper DETC2005-84631.
- [5] Zandarini, A., 2005. "Dynamic behaviour caracterization of a brake disc by means of electronic speckle pattern interferometry measurements". In Proceedings of the IDETC/CIE ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, Long Beach, California, USA, September 24-28, ASME, pp. 1–8. Paper DETC2005-84630.
- [6] Zandarini, A., 2007. "Full field espi measurements on a plate: challenging experimental modal analysis". In Proceedings of the XXV IMAC, Orlando (FL) USA, Feb 19-22., SEM. Paper s34p04.
- [7] Zandarini, 2009. "Full-field espi techniques in structural dynamics". In Atti del convegno Seconda giornata di studio Ettore Funaioli, 18 luglio 2008, Bologna, Italy, Maggiore, Meneghetti, and Parenti Castelli, eds., Quaderni di ricerca del DIEM, AMS Acta - Alm@DL Università di Bologna, pp. 31–56. See also URL <http://amsacta.cib.unibo.it/archive/00002552/>.
- [8] Zandarini, A., 2008. "Fatigue life assessment by means of full field espi vibration measurements". In Proceedings of the ISMA2008 Conference, September 15-17, Leuven (Belgium), P. Sas, ed., KUL, pp. 1–15. Paper 326.
- [9] Zandarini, A., 2008. "Full field espi vibration measurements to predict fatigue behaviour". In Proceedings of the IMECE2008 ASME International Mechanical Engineering Congress and Exposition, October 31- November 6, Boston (MA) USA, ASME, pp. 1–10. Paper imece2008-68727.
- [10] Zandarini, A., 2008. "Esp measurements in structural dynamics: fatigue life assessment". In Proceedings of the Dantec Dynamics 15th International Conference and User Meeting, Dantec dynamics GmbH, pp. 1–20.
- [11] Ewins, D. J., 2000. *Modal Testing - theory, practice and application*, 2nd ed. Research Studies Press Ltd., Baldock, Hertfordshire, England. ISBN 978-0-86380-218-8.
- [12] Heylen, W., Lammens, S., and Sas, P., 1998. *Modal Analysis Theory and Testing*. Katholieke Universiteit Leuven, Leuven (Belgium). ISBN 90-73802-61-X.

ANALYSIS OF THE EVOLUTION OF THE PRESSURE FORCES IN VARIABLE DISPLACEMENT VANE PUMPS USING DIFFERENT APPROACHES

Marco Cavallari

Engineering Department in Ferrara (EnDIF),

University of Ferrara, Italy

E-mail: marco.cavallari@unife.it

Alessio Montanari

Engineering Department in Ferrara (EnDIF),

University of Ferrara, Italy

E-mail: alessio.montanari@student.unife.it

Emiliano Mucchi

Engineering Department in Ferrara (EnDIF),

University of Ferrara, Italy

E-mail: emiliano.mucchi@unife.it

Giorgio Dalpiaz

Engineering Department in Ferrara (EnDIF),

University of Ferrara, Italy

E-mail: giorgio.dalpiaz@unife.it

Abstract. *Two models for the estimation of the pressure evolution in high pressure variable displacement vane pumps are proposed: a lumped parameter model and an empirical model. The former has full physical meaning and allows to simulate the behavior of a wide pump range, the latter, based on experimental measurements, can be applied only to a single family of pumps. Both the models can simulate the pressure evolution around the rotor in working conditions and the variable forces acting on the pump body as well. These results are important for the structural design of the pump. Moreover, these models are the first element of a combined model aimed at the dynamic simulation of this type of pump, as a tool for vibro–acoustical optimization.*

Keywords: *lumped parameter model, vane pump, pressure evolution, variable forces, noise and vibration*

1. INTRODUCTION

This paper deals with high pressure variable displacement vane pumps and presents a work carried out by the *Dept. of Engineering of the University of Ferrara*, in co-operation with

BERARMA s.r.l. (Casalecchio di Reno, Bologna, Italy).

This kind of volumetric pumps is widely used in machine tools and in hydraulic systems thanks to their control strategy. In fact the oil flow rate can be simply tuned accordingly to system and efficiency requirements. As a machine tool component, the pump being studied must provide low vibration and acoustic levels to ensure high working precision as well as health and comfort of technicians. In this sight, an optimization process must be performed to reduce vibration levels and noise emissions. It is well known that such a noise and vibration optimization can be a very expensive task, thus the use of models is suitable in order to simulate the dynamic behavior. In this context, this paper deals with different approaches for the simulation of pressure evolution and pressure forces in variable displacement vane pumps. In particular, two models are presented: firstly a *Lumped Parameter model (LP model)*, able to give answers to complex design task, and secondly a simpler *empirical model*, able to give a fast solution for easier design problems.

One of the problems a designer must solve in the frame of a noise and vibration optimization of such components is the reduction of pressure ripple, crucially important because the pressure ripple generates variable forces and body vibrations. An attempt to minimize the ripple phenomena has been described by Hattori et al. [1] and a prediction of the noise generated by a vane pump has been formulated in [2]. In a more general way, the proposed lumped parameter model uses an approach similar to [3–8], that are taken as a starting point for the development of this model.

Another important step of the present work is the validation procedure based on the comparison with the pressure evolution measured in working conditions. A work dealing with this topic has been carried out by Bianchini et al. [9].

The goal of these models is the estimation of the variable forces acting on the pump casing. In this light a former work has been performed by Novi et al. [10] to study the wear phenomena inside the pump.

As a matter of fact, in this type of pumps the variable pressure forces are the main sources exciting casing vibration; these forces can be estimated through the presented models and the reduction of the amount and rate of force variation can generally involve a reduction of casing vibration and noise emissions. With respect to previous works in the field of pump modeling, here the estimation of the pressure evolution and the pressure forces is the first step in the perspective of developing a combined model for the simulation of casing vibrations as well as noise propagation.

In more detail, further work will be needed in order to combine the pressure excitation forces and the structural frequency response functions of the pump casing. In this way the operational acceleration of the casing will be obtained and used as input data for a noise propagation model.

The organization of this paper is as follows. Section 2 presents the description of the variable displacement vane pump; in Section 3 the procedure for the experimental measurement of the pressure evolution is reported as well as the experimental results. Section 4 deals with the lumped parameter model while Section 5 describes the empirical model; for both models the pressure evolution estimation is presented and discussed. Section 6 then reports the results of the force calculation for both models and finally Section 7 is dedicated to the concluding remarks.

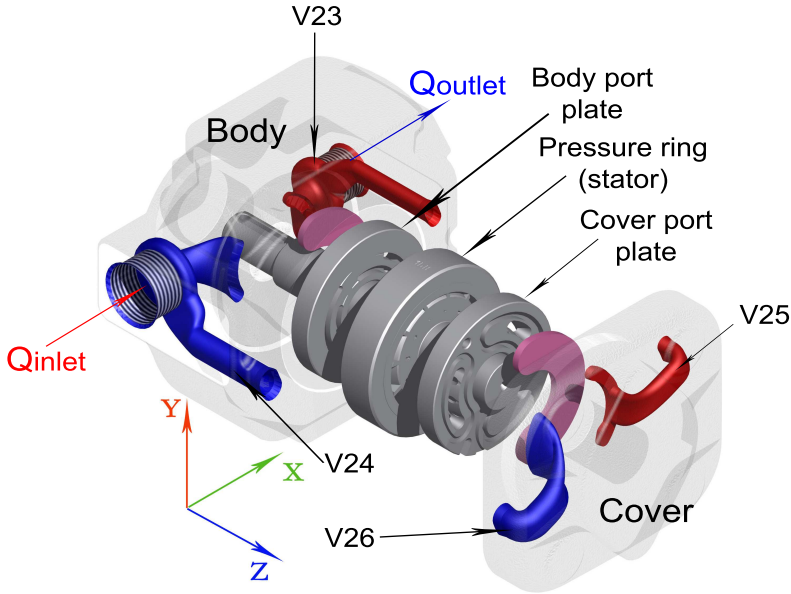


Figure 1: Exploded view of pump *PHV05* with distribution ducts.

2. VARIABLE DISPLACEMENT VANE PUMP DESCRIPTION

This work deals with high pressure vane pumps of the PHV series by *BERARMA s.r.l.* This kind of pumps can work in a pressure range from 20 to 250 bar. In particular *TYPE 01 PHV 05*, depicted in Fig. 1 and Fig. 2, and *TYPE 01PHV 1* have been studied. The first has an actual displacement of 16 cm^3 , the former of 32 cm^3 , both take advantage of a double inlet/outlet port plate system in which each port is packed on the pressure ring by using a system based on hydrostatic compensation.

The variable displacement vane pumps are a particular kind of rotating volumetric machines widely used in hydraulic applications. In fact they allow to regulate the lubricant flow rate, optimizing energy consumption. The variable displacement vane pump being studied is composed of a cast iron body, in which a one piece rotor shaft, a pressure ring and two port plates are located (see Fig. 1). The rotor presents radial grooves in which the vanes can slide. Hereafter, the volume delimited by two consecutive vanes, the pressure ring inner race, the rotor and the port plates is called vane space, while the volume defined by the rotor groove, the bottom of a vane and the port plates is defined as hole. The port plates present the outlet and inlet ports, suitably designed to communicate both with the vane spaces and the holes.

The pump can change its displacement in working conditions by varying the eccentricity between the rotor shaft and the pressure ring as it can be seen in Fig. 3. This operation is carried out by a control piston that moves the pressure ring, and by a bias piston that reports the pressure ring in the zero eccentricity condition when the desired pressure level is reached. The balancing screw depicted in Fig. 2 is used to compensate the resulting pressure force (see Section 6). This kind of pump can elaborate a full flow rate (see Fig. 3a) till the desired pressure level is reached, or it can work in a "zero flow condition" (see Fig. 3b), maintaining

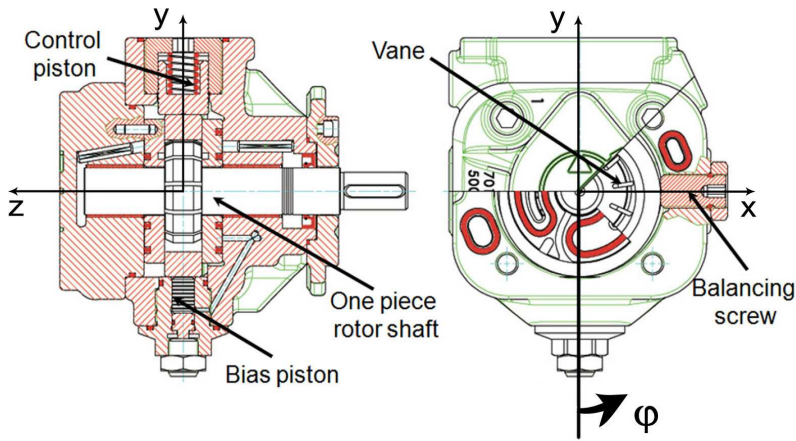


Figure 2: Variable displacement vane pump *PHV05*.

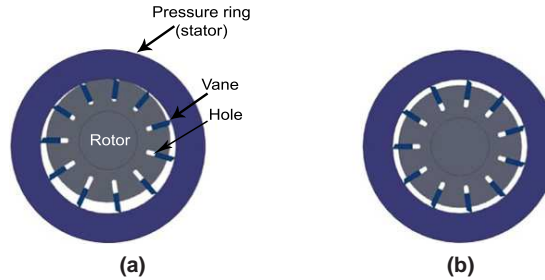


Figure 3: The two extreme working conditions: full flow (a) and zero flow (b) condition.

a desired pressure and compensating hydraulic losses.

The coordinate system of Fig. 1 will be used to calculate the variable forces in Section 6 while the pressure evolution is referred to φ coordinate, taken as positive in anti-clockwise sense (see Figs. 2 and 7).

3. PRESSURE EVOLUTION MEASUREMENTS IN WORKING CONDITIONS

Experimental tests on the pump have been carried out to validate the model results concerning pressure evolution and variable forces. The experimental rig (see Fig. 4) allows to measure the pressure inside a vane space and inside a hole for a complete rotation of the shaft by two piezoelectric pressure transducers directly placed on the one piece rotor shaft. To do this, a suitably designed rotor shaft is used (see Fig. 5) where two meatus connect a vane space and a hole to the end of the shaft, where the transducers are placed. The pump cover must be modified as well, to allow the rotor shaft passage and finally, two measure chambers (one containing the lubricant at the vane space pressure and one at the hole pressure) are isolated by using O-ring gaskets. The pressure signals and the trigger are acquired and processed by *LMS Scadas SC305* front end controlled by *LMS Test.Lab* software. The pressure sensor

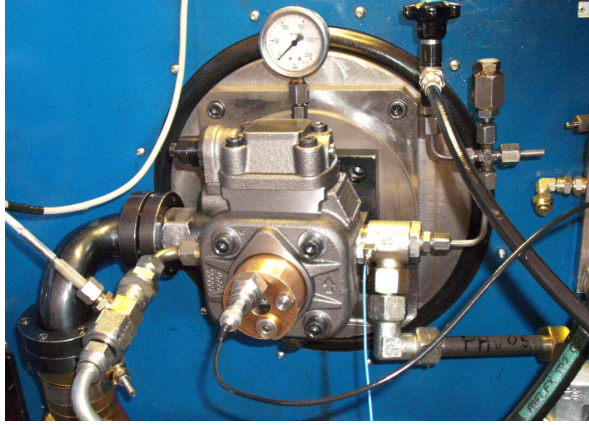


Figure 4: Pump *PHV05* on the test bench.

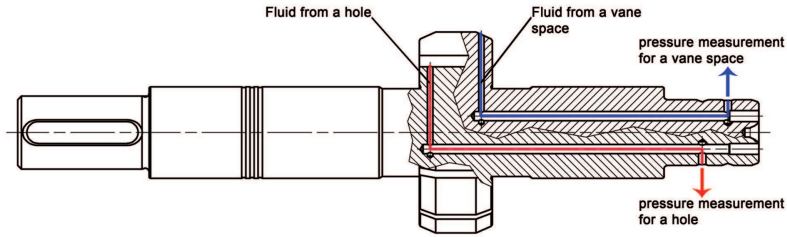


Figure 5: The modified rotor shaft used for the tests.

being used is a piezoelectric transducer *Kistler (model 6207)*, the proximity sensor used as trigger is a *Balluff BES 516324S4C*.

All the results presented in this paper – both experimental and simulation results – concern operational pressure of 250 bar and rotational speed of 1500 rpm.

With this set up the measure of the pressure inside a hole and a vane space can be obtained. The results are shown in Fig. 6, in which the pressure evolution is presented for a complete shaft rotation. More precisely, it deals with the pressure evolution in a vane space during the shaft rotation; the initial position of the pressure evolution corresponds to the position shown in Fig. 7 for vane space 1 and hole 12. The passage from the inlet to the outlet region can be immediately distinguished and the ripple phenomena for the high pressure region are captured as well.

4. THE LUMPED-PARAMETER MODEL

General description

Several researches have developed Lumped Parameter models (*LP models*) for vane pumps in order to obtain the pressure evolution around the rotor shaft. The lumped parameter approach used in this work is based on [11, 12]. These *LP models* are based on a Euler's approach

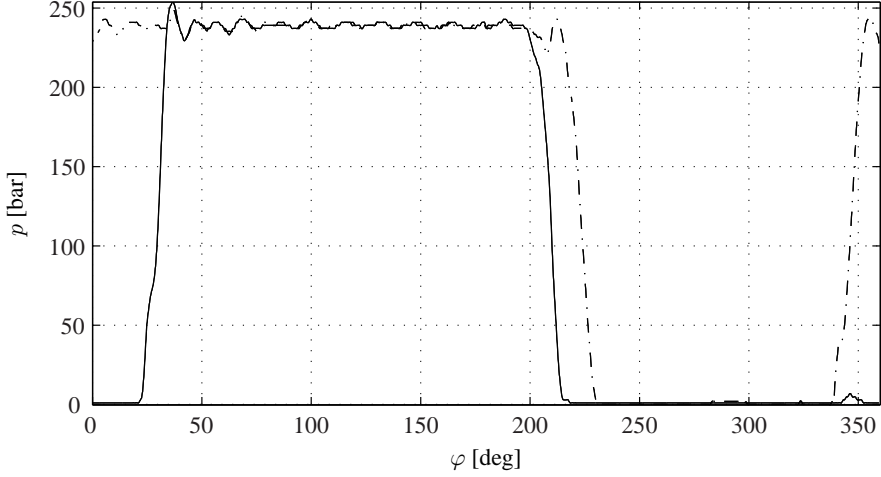


Figure 6: Pressure evolution measured in a vane space (solid line) and in a hole (dash-dot line) for a complete shaft rotation (pump *PHV05*). Full flow condition.

in which the pressure evolution during a rotation is studied by dividing the pump in several control volumes corresponding to vane spaces; nevertheless they were not able to capture the pressure ripple. Subsequent works can be found in the literature improving this first approach by increasing the number of control volumes to appreciate the pressure ripple in a vane space [5–8]. In these last models also the distribution ducts are treated as control volumes. The present work is based on the latter approach but accounts also the pressure calculation in the holes. The pressure in this domain is very important in order to define the dynamical behavior of the rotor, but also to calculate the reaction forces between the vanes and the pressure ring.

The first step in developing a lumped parameter model is the assumption of proper coordinate systems, as shown in Fig. 1 and Fig. 2. Then, the pump system must be divided in several control volumes. The present model is based on 26 control volumes: control volumes from V_1 to V_{11} correspond to the vane spaces; control volumes from V_{12} to V_{22} correspond to holes, control volumes V_{23} and V_{25} correspond to the outlet ducts and control volumes V_{24} and V_{26} to the inlet ducts (see Figs. 1 and 7). In particular, control volume V_{23} (V_{24}) is the outlet (inlet) control volume that directly communicates with the hydraulic system, while V_{25} (V_{26}) is not directly linked with the plant, but through V_{23} (V_{24}), as depicted in Fig. 1.

It is now possible to apply the flow rate continuity equation to generic control volume i : each control volume is treated as an open thermodynamic system with mass transfer with its surroundings. The rate of change in fluid pressure induced by the mass transfer and volume variation, being the fluid characterized by its isothermal bulk modulus B_{oil} , can be conveniently expressed in terms of rotor shaft angular coordinate φ , positive in the anti-clockwise (see Fig. 2) [3]:

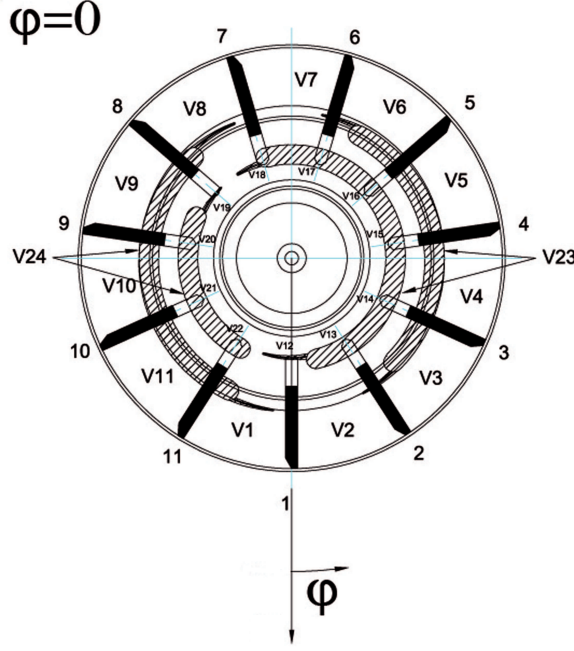


Figure 7: Control volumes on the xy plane.

$$\frac{dp_i}{d\varphi} = \frac{B_{oil}}{V_i \omega} \left(\Delta Q_i - \frac{dV_i}{d\varphi} \omega \right) \quad (1)$$

By using this relation, it is possible to determine the pressure variation inside control volume i caused by flow rate gain ΔQ_i and by the volume variation of $\frac{dV_i}{d\varphi}$. In Eqn. (1), ω is the rotor angular speed, p_i represents the pressure in control volume i , V_i is its volume and flow rate gain ΔQ_i is the balance between the flows coming in and out from control volume i . Flow rates of different types are properly taken into account, depending on the considered control volume, as indicated below: laminar, drag, and turbulent flows (see Figs. 8a and 8b). The laminar flow rates (see Fig. 8b) take place in the vane flank meatus, from a vane space to the adjacent vane spaces (Q_{vf}), from a vane space to the holes (Q_{vh}), in the meatus between pressure ring and rotor (drainage flow rate Q_{vs}) and in the meatus between the rotor and the port plates (drainage flow rate Q_{vd} for the vane spaces, Q_{hd} for the holes). Since the vane is supposed always in contact with the inner race of the pressure ring, the flow rate in correspondence of the vane head is not taken into account. All these flow rates can be calculated as suggested by Eqn. (2), in which $b_i(\varphi)$, $h_i(\varphi)$ and $l_i(\varphi)$ are respectively the meatus width, thickness and length, quantity μ represents the lubricant dynamic viscosity, p_i is the pressure in the considered control volume and p_j is the pressure in the control volume interested by the exchange:

$$Q_{lam,i} = \frac{b_i(\varphi)h_i(\varphi)^3}{12\mu l_i(\varphi)} (p_j(\varphi) - p_i(\varphi)) \quad (2)$$

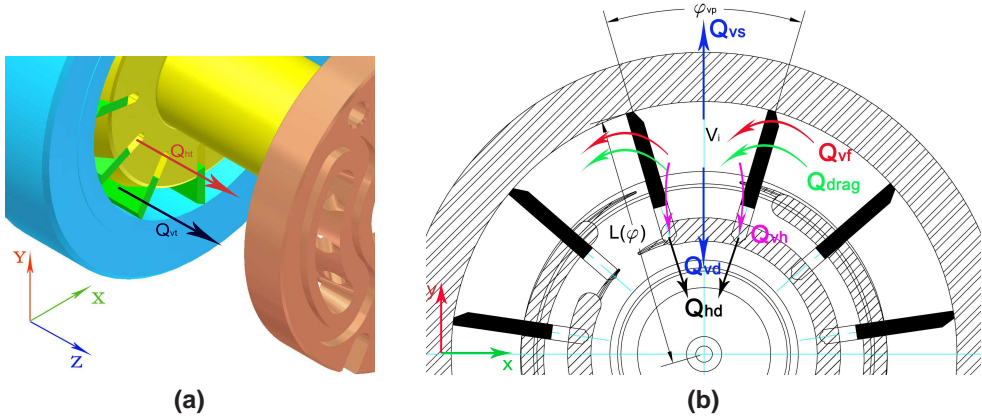


Figure 8: Flow rates involved in the model: flows in z direction (a) and in the xy plane (b).

The drag flow interests only the vane spaces and it is not dependent on the pressure field. It can be calculated as depicted in Eqn. (3), in which quantity $h_{f,i}$ represents the clearance between the vane flank and the distributor, $L_i(\varphi)$ is the distance between the rotor center and the pressure ring at angular coordinate φ and $L_i(\varphi - \varphi_{vp})$ represents the distance between the rotor center and the pressure ring at the coordinate $\varphi - \varphi_{vp}$, where φ_{vp} is the attitude angle by a vane pitch (see Fig. 8b):

$$Q_{drag,i} = h_{f,i} \frac{\omega}{2} (L_i(\varphi)^2 - L_i(\varphi - \varphi_{vp})^2) \quad (3)$$

The turbulent flow rates take place between the vane spaces and the distribution ducts (Q_{vt}), between the holes and the distribution ducts (Q_{ht}) and between the ducts and the hydraulic system (Q_{inlet} , Q_{outlet}) as shown in Figs. 1 and 8a. They can be calculated using Eqn. (4) in which k represents the flow coefficient, $A_i(\varphi)$ is the exchange area from control volume i to control volume j , ρ is the lubricant density, p_i is the pressure in the considered control volume and p_j is the pressure in the control volume interested by the exchange:

$$Q_{turb,i} = k A_i(\varphi) \sqrt{\frac{2|p_j(\varphi) - p_i(\varphi)|}{\rho}} \text{sign}(p_j(\varphi) - p_i(\varphi)) \quad (4)$$

It is now clear that quantity ΔQ_i of Eqn. (1) depends on the pressure field so, applying Eqn. (1) to all the control volumes, a system of 26 differential equations (5) can be obtained and the unknown quantities, i.e. the pressure evolution in each control volume can be calculated:

$$\left\{ \begin{array}{l} \frac{dp_1(\varphi)}{d\varphi} = f_1(p_1, \dots, p_{26}) \\ \vdots \\ \frac{dp_{26}(\varphi)}{d\varphi} = f_{26}(p_1, \dots, p_{26}) \end{array} \right. \quad (5)$$

where functions $f_1 \dots f_{26}$ define the right side of continuity equation (1).

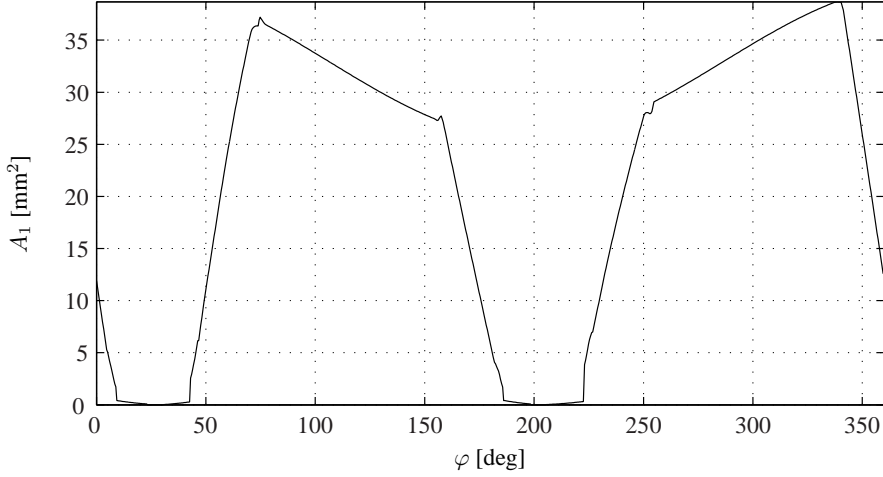


Figure 9: Exchange area between the vane space corresponding to control volume 1 and the distribution ducts (pump *PHV05*).

Before integrating equation system (5), some auxiliary quantities must be calculated. For example, with reference to Eqn. (1), volume V_i and volume variation $\frac{dV_i}{d\varphi}$ must be calculated for each control volume. Looking at Eqn. (2), quantities $b_i(\varphi)$, $h_i(\varphi)$, $l_i(\varphi)$ must be defined on the basis of the clearances inside the pump, as well as quantity $L_i(\varphi)$ in Eqn. (3). Finally in Eqn. (4), exchange area $A_i(\varphi)$ has to be fully defined for the volumes of the vane spaces and of the holes. All these preliminary calculation are solved by using subroutines in MATLAB. As an example the exchange area for a vane space as a function of the angular coordinate φ is depicted in Fig. 9. This quantity varies with angular coordinate φ but it is not influenced by the pressure ring eccentricity. Other quantities, such as volume V_i takes into account also such an influence by including the value of the eccentricity imposed by the control piston together with the eccentricity due to the rotor displacement in working conditions.

To take into account the effect of the plant, the outlet flow rate Q_{outlet} (see Fig. 1) has been expressed as follows:

$$Q_{outlet} = k_{outlet} A_{outlet} \sqrt{\frac{2|p_{atm} - p_{23}(\varphi)|}{\rho}} \text{sign}(p_{atm} - p_{23}(\varphi)) \quad (6)$$

Equation (6) refers to the outlet duct control volume 23 which is the only outlet control volume that directly communicate with the hydraulic circuit. In Eqn. (6), quantity p_{atm} represents the constant atmospheric pressure and coefficient k_{outlet} takes into account the plant losses and resistance. By using Eqn. (6) it is possible to estimate flow coefficient k_{outlet} ; in fact, assuming the experimental outlet flow rate Q_{exp} as Q_{outlet} , the pressure in the outlet chamber as p_{outlet} (constant value, equal to operational pressure). Coefficient k_{outlet} is equal to 0.0058 in the full flow condition and $4 \cdot 10^{-9}$ in the zero flow condition, calculated as follows:

Table 1: Main parameters of pump *PHV05*.

B_{oil}	Lubricant bulk modulus	$1.72 \cdot 10^9$ Pa
μ	Lubricant dynamic viscosity	0.014 Pa s
ρ	Lubricant density in std conditions	854 kg/m ³
φ_{vp}	Vane pitch	32.73 deg
N	Number of vanes	11
W_s	Pressure ring width	20 mm
r_s	Pressure ring inner radius	30 mm
r_r	Rotor shaft radius	25 mm
t_{vb}	Vane thickness at the base	2.2 mm
t_{vh}	Vane thickness at the head	0.3 mm
v_m	Vane mass	0.003 kg
d_w	Inlet and outlet ducts width	3.5 mm

$$k_{outlet} = \frac{Q_{exp}}{A_{outlet}} \sqrt{\frac{\rho}{2|p_{atm} - p_{outlet}|}} \quad (7)$$

Finally, once all the quantities underlined in the present section are preliminary calculated using as main input data the values contained in Tab. 1, it is possible to launch the integration iterative procedure for equation system (5). Since all the phenomena involved in the pump rotation are periodic with vane pitch φ_{vp} , at constant operational speed the integration is done in the angular domain, from $\varphi = 0$ to $\varphi = \varphi_{vp}$.

An initial value of the pressure in the control volumes is required for each control volume, then an iterative procedure can start by using as initial condition for the subsequent iteration the results of the previous iteration and after several cycles the solution for the pressure evolution is obtained. The integration procedure stops once the difference between the pressure forces calculated after two subsequent iterations is under a tolerance value a priori established. Then, the pressure evolution around the shaft is obtained by adding the pressure field of each control volume consecutively.

Results and validation

Figures 10a and 10b depict the results in terms of pressure evolution for a complete shaft rotation in a vane space and in a hole respectively. Quantity p^* represents the normalized pressure (see Eqn. (8), p_{outlet} is the constant operational pressure in the outlet chamber):

$$p^* = \frac{p}{p_{outlet}} \quad (8)$$

Both the plots depict pump behavior in the full flow rate condition. The pressure evolution are compared and validated with respect to measured data (see Fig. 6) by using *Percentage of Good points* index (*PGP*) [13]. Such an index assumes higher values when the curves tend to coincide. In practice, a tolerance range δ has to be defined a priori (in this case δ is equal to 3%). Consequently, when a good resemblance between the measured data and the simulated one is obtained, a large part of the points of the latter will lie inside the tolerance area and the *PGP* will assume values tending to 100%. The *PGP* for the pressure evolution in a vane space and in a hole are 71% for both the full flow and zero flow condition.

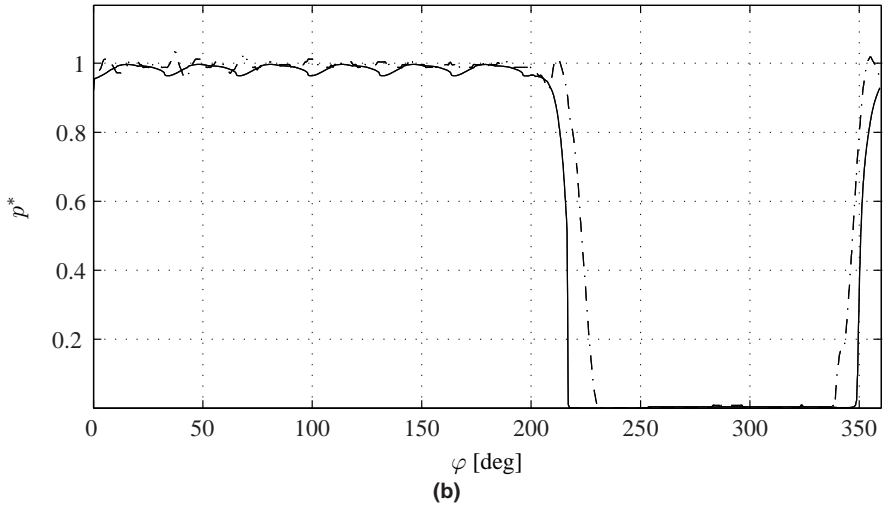
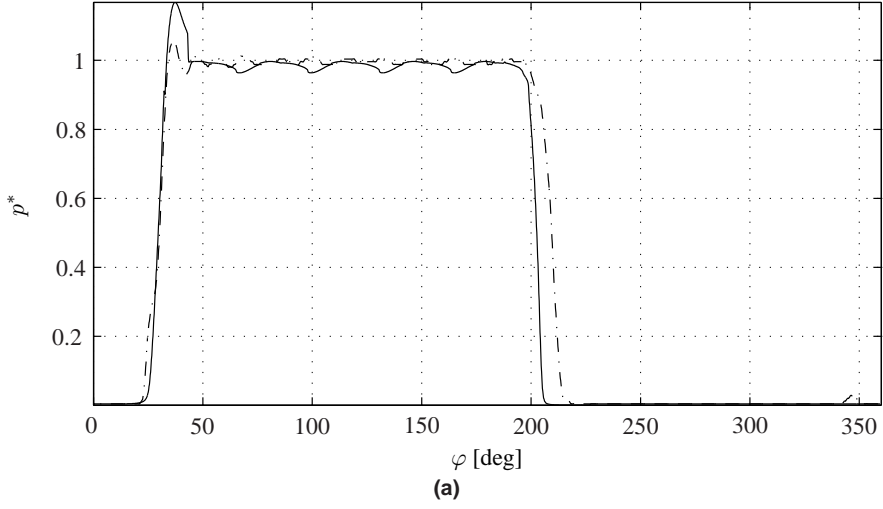


Figure 10: Normalized pressure evolution for a complete shaft rotation in a vane space (a) and in a hole (b) referred to pump *PHV05*: experimental (dash-dot line) and *LP* model results (solid line). Full flow condition.

The simulated pressure shows a good concordance with the measured pressure and the *PGP* index gives evidence to this fact. Figs. 10a and 10b show that the simulated pressure evolutions capture the ripple phenomena. However a few mismatches between the experimental quantities and the simulated ones can be noticed, as the ripple shape, the peak amplitude in a vane space at the beginning of the high pressure zone and the instant when the pressure increases and decreases.

5. THE EMPIRICAL MODEL

General description

The empirical model has been developed in order to predict the pressure evolution inside the pump by using an approach focused on industrial applications. In fact this model is fast to develop and easy to use. It is based on data obtained with the experimental measurements described in section 3 and for this reason its application should be limited to pumps belonging to families similar to those used during the test campaign.

The experimental pressure evolution analysis performed on measured data from 50 up to 250 bar shows that the pressure trend strongly depends on geometrical parameters (displacement, port plate design) and operational parameters (shaft velocity and working pressure). In more detail, pressure evolution curve $p(\varphi)$ can be subdivided into four regions (see Fig. 11):

- A region characterized by the inlet (atmospheric) pressure: before A and after D ;
- A region where the pressure increases from the inlet to the outlet pressure: from A to B ;
- A region at about the outlet pressure: from B to C ;
- A region where the pressure decreases from the outlet to inlet pressure: from C to D .

For the *PHV* series, several experimental pressure evolutions were analysed, obtained at different values of outlet pressure, shaft velocity, displacement, number of vanes and port plate geometry, both at the zero flow and full flow condition. By using all these experimental plots as a data-base, the trend of the above-mentioned regions has been correlated to the operational and geometrical pump parameters, enabling to synthesize the analytical equation of each region of the pressure evolution curve. The first region (before A and after D) of the synthesized pressure curve has been always imposed to be equal to the atmospheric pressure, as in the measured curves. The second region consists in a transition from the inlet to the outlet pressure, starting and ending in defined angular coordinates that mainly depend on the port plate design. The peak at the end of this transition (Δp_{peak} , see Fig. 11) is strongly influenced by the outlet pressure and the operational condition: the peak assumes larger amplitude in the full flow condition than in the zero flow condition. Furthermore, the higher the displacement is, the largest the peak amplitude becomes. Therefore, this region has been synthesized with a cubic transition from the inlet to the outlet pressure, where the starting and ending positions (φ_1 and $\varphi_1 + \Delta\varphi$ respectively) are governed by the port plate design and the peak amplitude is influenced by the operational condition, outlet pressure and displacement, as in the experimental data.

The fourth region has been synthesized at the same manner as the second region, since they depend on the same pump parameters.

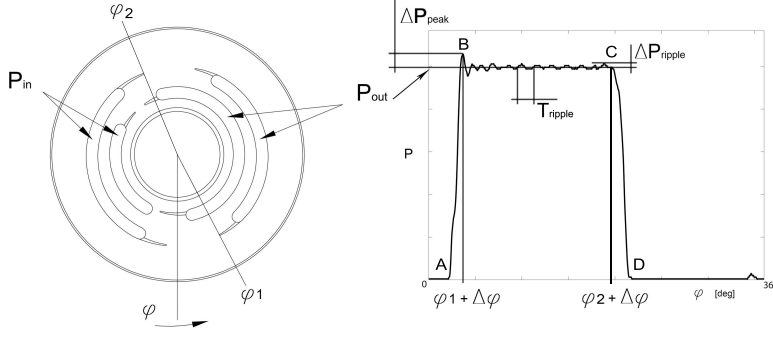


Figure 11: Empirical model: pressure evolution parameters.

In the experimental pressure evolution plot, the third region presents the genuine ripple phenomenon around the outlet pressure. The amplitude of the ripple (Δp_{ripple}) has been correlated to the operational condition, displacement and outlet pressure. With the aim of detecting the periodicity of this ripple (T_{ripple}), an order analysis has been performed in the third region leading to the spectrum of Fig. 12, where the first order corresponds to the frequency equal to the shaft rotational frequency (the mean component has been neglected). As expected, the main order components occur in correspondence of the 11th, 22th and 33th harmonics, since the pump being studied has 11 vanes. Taking advantage of this analysis, this third region has been synthesized as a Fourier series, considering only such three high amplitude order harmonics (vane harmonics):

$$p(t) = p_{outlet} + \sum_{k=1}^3 a_k \cos(kN\omega t + \varphi_k) \quad (9)$$

where the amplitude and phase of each term of the Fourier series (a_k , φ_k) have been correlated to a few pump parameters (operational condition, displacement, outlet pressure) while the periodicity depends on the vane number (N) and on the rotor angular speed (ω).

In this sight, the pressure evolution for a complete shaft rotation can be analytically synthesized as a function of the main pump operational and geometrical parameters. As a summary, Tab. 2 collects in the first column the parameters used for the pressure evolution synthesis and in the second column the relative physical pump parameters.

Once the algorithm for the estimation of the pressure evolution has been developed, it can be used for the estimation of the pressure evolution of new prototypes of pumps with different displacements, belonging to families similar to those used during the test campaign.

Results and prediction

The model results have been compared with the experimental tests results by using *PGP* index (see Section 4). A good concordance between predictions and the measured data is shown. In Figs. 13 and 14 the results for *PHV05* pump with an actual displacement of 16 cm³ and *PHV1* with an actual displacement of 32 cm³ are shown. In Tab. 3 the results

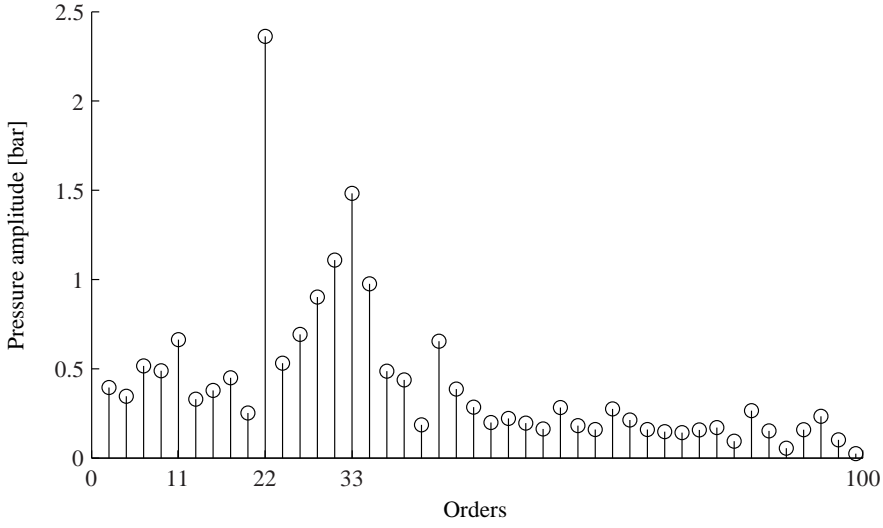


Figure 12: Experimental pressure ripple spectrum (pump PHV 05, vane space, full flow condition) computed from the pressure evolution of Fig. 6 in the range from 50 deg to 180 deg.

Table 2: Pressure evolution parameters and correlations to pump parameters.

Pressure evolution parameters	Pump parameters
$\Delta\varphi, \varphi_1, \varphi_2$	Port plate geometry
Δp_{peak}	Operational condition and pressure, displacement
Δp_{ripple}	Operational condition and pressure, displacement
T_{ripple}	Number of vanes

for *PGP* index are listed: the models offers good results both in full flow and in zero flow condition.

Since the model gave good results for the pumps whose experimental measurements have been used as "data-base", it has been used for the prediction of the pressure evolution of pump *Type 01 PHV3*. Such a pump belongs to a family similar to those used during the test campaign. The estimation depicted in Fig. 15 shows how the increment in displacement mainly affects the amplitude of Δp_{peak} , dramatically increased with respect to the peak of the smaller pumps. Finally, the experimental evidence and model results show how the geometry, with reference to the port plate ducts, has a great influence on the pressure distribution, in particular on the transition regions (see φ_1 and φ_2 in Fig. 11). Going further, a proper design of the carvings at the beginning and at the end of the ducts can limit strong pressure variation in the transition regions, especially at high working pressure.

For the sake of completeness, the comparison of results obtained in the tests, by the *LP* model and by the empirical model at zero flow condition for a vane space and for a hole of *PHV05* pump is reported in Fig. 16 and the good concordance between the three curves is quantitatively defined by the *PGP* indexes resumed in Tab. 3.

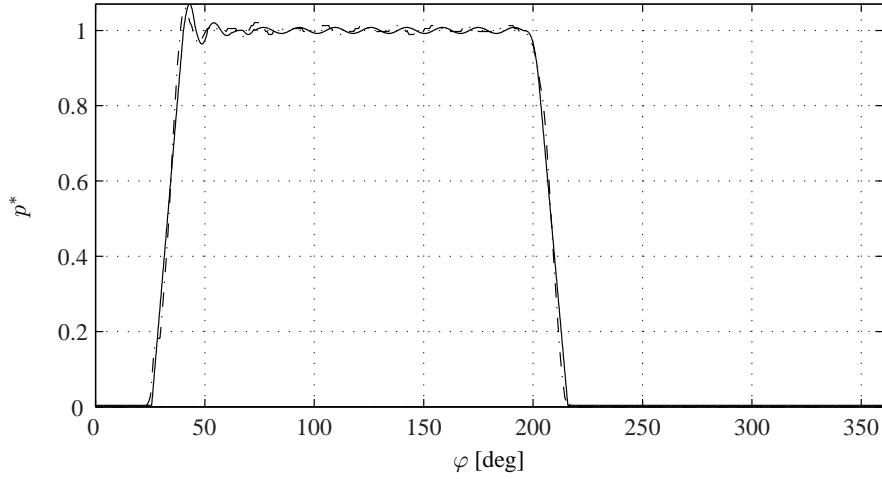


Figure 13: Normalized pressure in a vane space predicted by the empirical model (solid line) and measured pressure in working conditions (dash-dot line) for the pump *PHV05*. Full flow condition.

Table 3: *PGP* index with tolerance range of 3% comparing the simulated and measured pressure evolutions.

	<i>PHV05</i>		<i>PHV1</i>
	<i>LP</i> model <i>vs</i> test	Empirical model <i>vs</i> test	Empirical model <i>vs</i> test
Full flow	71%	87%	86%
Zero flow	71%	77%	74%

Some general observation about pressure evolutions can be finally done. It is possible to notice that the pressure starts to increase when the communication between inlet ducts and a vane space or a hole is interrupted (i.e. at 22.5 deg for the *PHV05* pressure evolution in a vane space). Hence, a peak is clearly visible in the vane space pressure evolution in full flow condition (see Fig. 10a) once the outlet distribution groove (visible before the outlet duct, in Fig. 7) is completely passed (40 deg for the *PHV05*); then the pressure evolution starts to decrease at the beginning of the inlet groove (200 deg for *PHV05*) to finally reach the inlet pressure value when the inlet duct fully communicates with the vane space or the hole (220 deg for the vane space pressure evolution in *PHV05* pump, see Fig. 10a).

6. DYNAMIC RADIAL FORCE

Force calculation

The pressure distributions inside the vane spaces and inside the holes are used to determine the resulting force acting in the radial direction (*xy* plane). This force is very important for the definition of the dynamic behavior of the pump casing from a vibrational and acoustical point

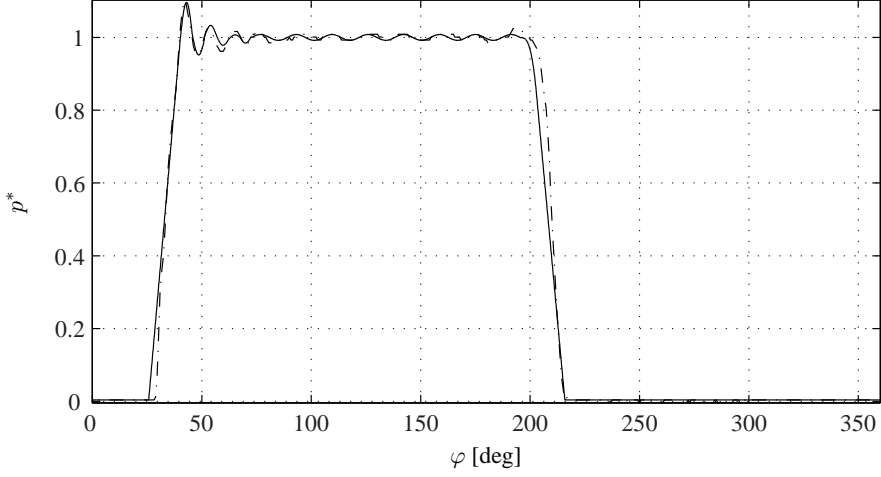


Figure 14: Normalized pressure in a vane space predicted by the empirical model (solid line) and measured pressure in working conditions (dash-dot line) for the pump *PHV1*. Full flow condition.

of view. In fact, these components are the forces that can excite casing vibration and produce noise. This force acting on the inner race of the pressure ring (F_{stator}) can be decomposed into two components (see Fig. 17):

- $F_{stator,x}$ along x direction, it loads the balancing screw;
- $F_{stator,y}$ along y direction, it loads the control piston.

Both the components are variable with respect to angular coordinate φ with period φ_{vp} ; for this reason all the results presented hereafter are plotted in this angular domain. In more detail, with reference to Fig. 18a, the pressure force acting on the pressure ring can be obtained by the contributions of the force due to pressure in vane space i , namely $F_{vs,i}$ and the contact force between the vane and the inner race of the pressure ring, $F_{v,i}$. These components must be summed for all the vane spaces. Equation (10), in which N represents the number of control volumes involved in the force calculation (equal to the number of vanes), summarizes this procedure. Term $F_{vs,i}$ can be approximately calculated as suggested in Eqn. (11), in which W_s represents the pressure ring width, \vec{n}_i defines the angle bisector of vane pitch i (see Fig. 18a) and the meaning of $L(\varphi)$ is shown in Fig. 8b and Section 3. The integration of Eqn. (11) takes place for a complete vane pitch (φ_{vp}). Term $F_{v,i}$ can be expanded as in Eqn. (12).

$$\vec{F}_{stator}(\varphi) = \sum_{i=1}^N \vec{F}_{vs,i}(\varphi) + \vec{F}_{v,i}(\varphi) \quad (10)$$

$$\vec{F}_{vs,i}(\varphi) = \left(W_s \int_{-\varphi_{vp}/2}^{\varphi_{vp}/2} p_i \cos \varphi L_i(\varphi) d\varphi \right) \vec{n}_i \quad (11)$$

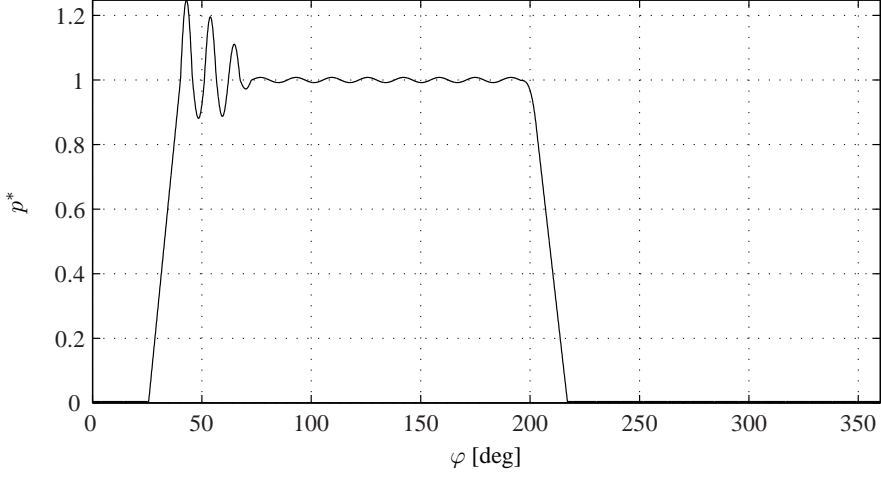


Figure 15: Normalized pressure in a vane space predicted by the empirical model for the pump *PHV3*. Full flow condition.

$$\vec{F}_{v,i}(\varphi) = \vec{F}_{hole,i}(\varphi) + \vec{F}_{centr,i}(\varphi) - \vec{F}_{head,i}(\varphi) \quad (12)$$

where $F_{hole,i}$ represents the force component due to the pressure in holes, $F_{centr,i}$ is the centrifugal force and $F_{head,i}$ is the component due to the pressure evolution in vane spaces (see Fig. 18b).

Results and comparison

As stated in Section 2, this kind of pumps can work at different flow conditions, from a "zero flow" condition to a "full flow" condition. In the previous Sections, the experimental and simulation results for the pressure evolution have been shown for both full flow and zero flow conditions. For the sake of brevity the results about pressure force will be presented hereafter only for the zero flow condition. The radial pressure force evolution acting on a vane pitch is depicted in Fig. 20, comparing the force calculated by using the measured pressure evolution, the pressure evolution derived by the empirical model and by the lumped parameter model.

The comparison between Fig. 20a and Fig. 20b shows that, nevertheless the force in y direction has lower mean value than in x direction, the most relevant variation is related to y direction; in particular, the *LP* model force exhibits a sudden variation of 1200 N at about 12 deg in y direction, while the corresponding sudden variation in x direction is about 400 N. This is important in the sight of noise and vibration optimization, since dynamic effects are mainly related to the amount and rate of force variation.

It is finally possible to identify the event related to such sudden force variations in Figs. 20a and 20b. In Fig. 16 the pressure evolutions for holes and vane spaces in zero flow condition are depicted: the x axis has been divided into 11 parts, each one corresponding to a vane pitch. It is possible to notice that a strong variation in the pressure field occurs for vane spaces in the first part of the 7th vane pitch (see Fig. 16a, from 196.3 deg to 229.1 deg).

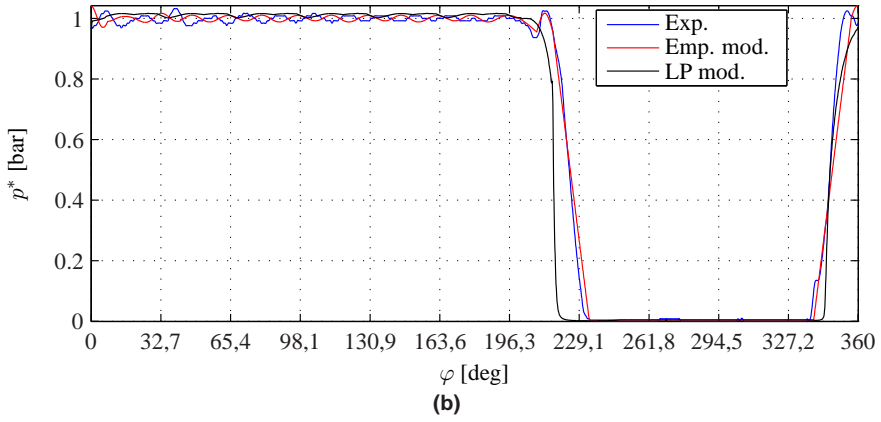
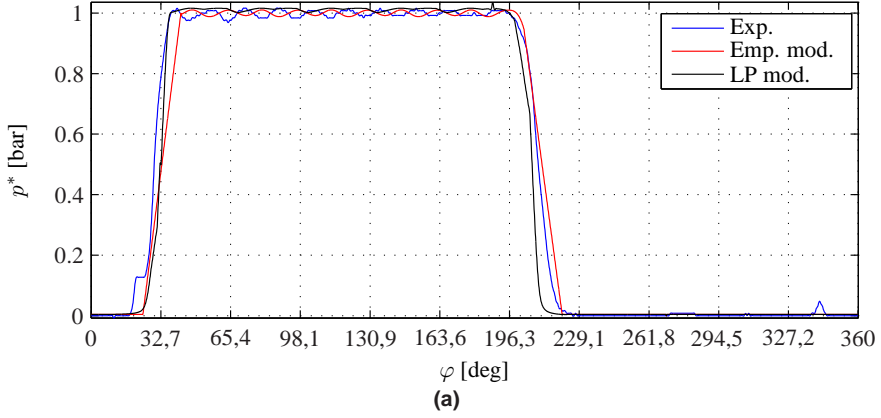


Figure 16: Comparison between measured pressures (blue line), pressures obtained by empirical model (red line) and pressures obtained by lumped parameter model (black line), for a vane space (a) and for a hole (b). Zero flow condition, *PHV05* pump.

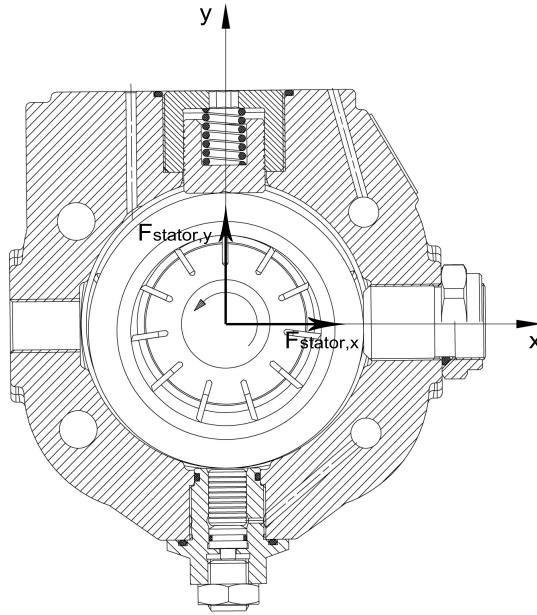


Figure 17: Radial pressure forces acting on the pump casing.

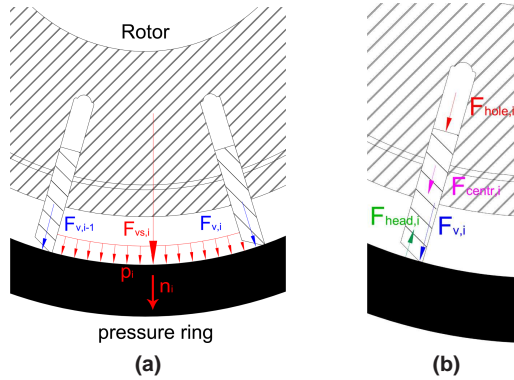


Figure 18: Radial forces acting on the pressure ring: contribution of the pressure evolution inside a vane space (a) and contribution of the vane head contact force (b).

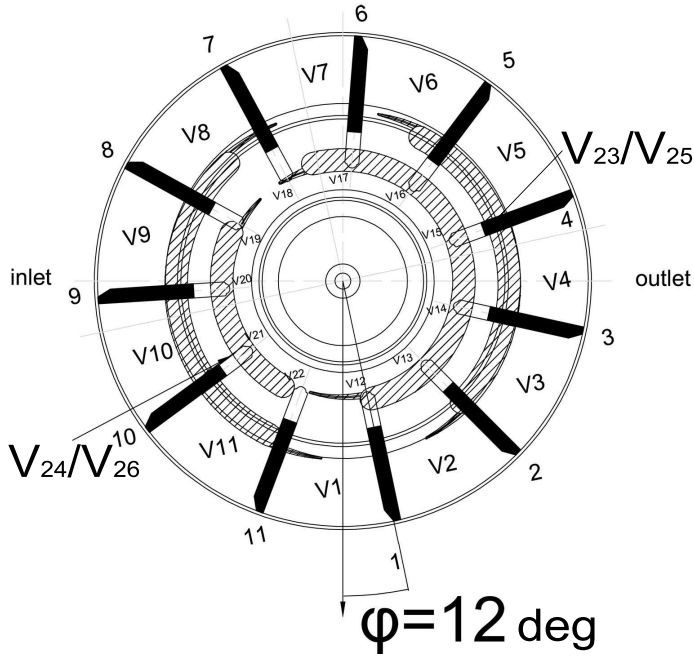


Figure 19: Position corresponding to 12 deg after the initial position.

Furthermore, pressure evolution simulated by the *LP* model shows more sudden changes with respect to the one predicted by the empirical model and the pressure evolution measured in working conditions, both for vane spaces and for holes. For this reason the force evolutions predicted by the *LP* model differ from the force evolutions predicted by the empirical model and the force evolution calculated from the pressure measured in working condition, especially around 12 deg, corresponding to the angular region previously identified. At this angular coordinate the vane space control volume V_7 and the hole control volume V_{19} start to communicate at the same time with inlet control volumes V_{24} and V_{26} (see Fig. 19). This fact involves the pressure drops depicted in Fig. 16, in correspondence of the 7th vane pitch and finally the force sudden variations depicted in Figs. 20a and 20b.

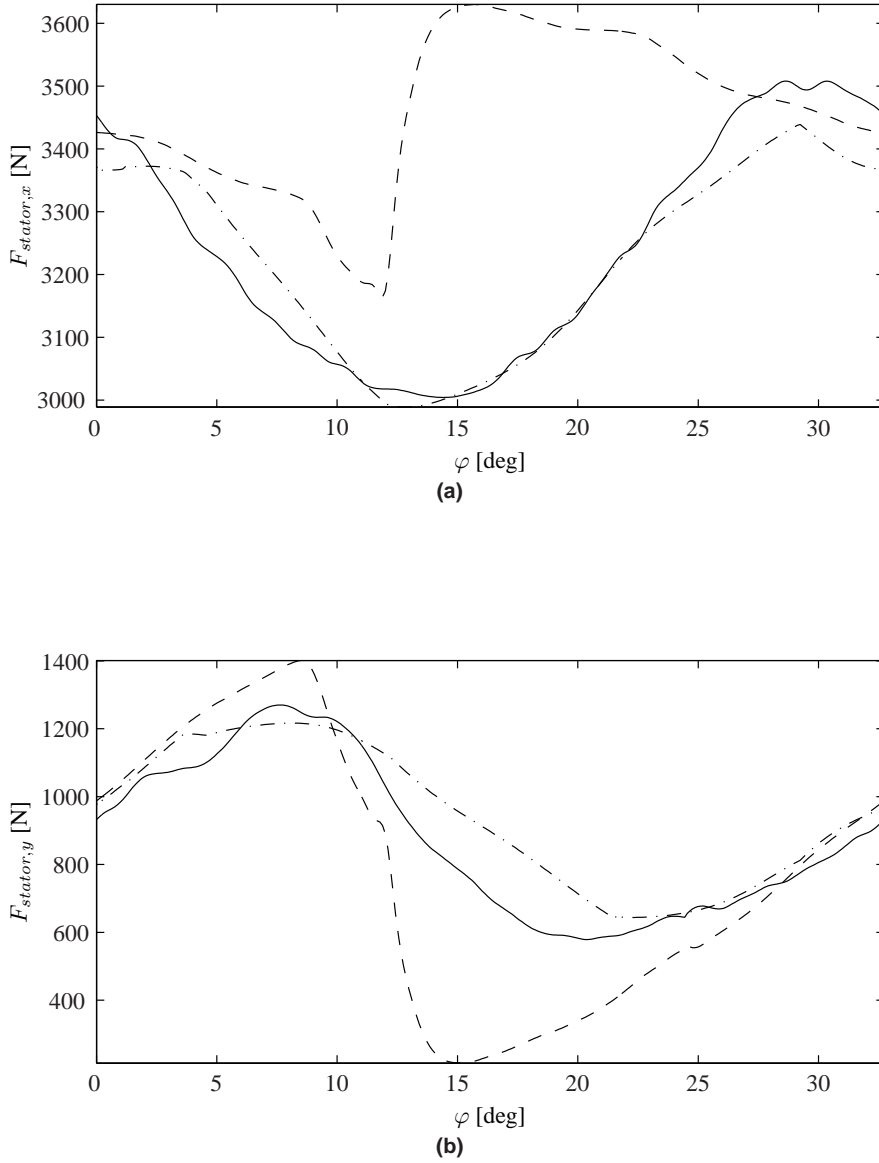


Figure 20: Radial pressure forces acting on the pressure ring over one vane pitch for the pump PHV05: forces from measured data (solid line), forces from the empirical model (dash-dot line), forces from the LP model (dashed line), at zero flow condition, along x direction (a) and y direction (b).

7. CONCLUDING REMARKS

This paper aims to describe an industrial research carried out in co-operation with *BERARMA s.r.l.* about their variable displacement vane pumps of the *PHV* series. The first target of this research is to develop a reliable tool for the designer who wants to investigate different solutions without employing long and expensive experimental tests. As a first answer an empirical model has been developed, based on a series of experimental measurements. This kind of model can give fast and good estimations of the pressure and force evolution only for a restricted family of pumps. At the same time a lumped parameter model has been developed whose analytical formulation allows to take into account a wider range of design and operational parameters. In terms of accuracy, the empirical model gives better results both in full flow and zero flow conditions: the *PGP* index for pressure evolution referred to the pump *PHV05* is 0.87 with respect to 0.71 obtained by the lumped parameter model in full flow condition and 0.77 with respect to 0.71 in zero flow condition. The main discrepancy in the pressure evolution between *LP* model and tests is related to the region where the pressure decreases from the outlet to the inlet pressure (see Figs. 10 and 16). Both the models allow to calculate the force distribution on the pressure ring matching the main periodicities. It has to be noticed that, while the empirical model gives a good estimation of forces, the lumped parameter model overestimates the force variation amplitude along the *y* direction and gives a force evolution in *x* direction characterized by a sudden variation not recognizable in the force evolution calculated from measured pressure data (Fig. 20). In this sight the *LP* model requires adjustments devoting at reducing the differences between simulated and measured quantities; however results in terms of forces can be considered promising and they can be used as input data for a noise and vibration optimization of the pump body.

Acknowledgements

This work has been developed within the Laboratory of Advanced Mechanics (MechLav) of Ferrara Technopole, realized through the contribution of Regione Emilia Romagna – Assessorato Attività Produttive, Sviluppo Economico, Piano telematico – *POR – FESR* 2007-2013, Activity I.1.1.

REFERENCES

- [1] Hattori, K., Suzuki, H., and Hasegawa, J., 1987. “Design method of a small-ripple vane pump”. *SAE International*(871681).
- [2] Dickinson, A. L., Edge, K. A., and Johnston, D. N., 1993. “Measurements and prediction of power steering vane pump fluidborne noise”. *Proceedings of the 1993 Noise and Vibration Conference*.
- [3] Mucchi, E., Dalpiaz, G., and Fernández del Rincón, A., 2010. “Elastodynamic analysis of a gear pump. part 1: Pressure distribution and gear eccentricity”. *Mechanical Systems and Signal Processing*, **24**, pp. 2160–2179.
- [4] Mucchi, E., 2007. “Dynamic analysis of external gear pumps by means of non linear models and experimental techniques”. PhD thesis, Università degli Studi di Ferrara - Engineering Department in Ferrara, March.
- [5] Barbarelli, S., Bova, S., and Piccione, R., 2009. “Zero-dimensional model and pressure data analysis of a variable-displacement lubricating vane pump”. *SAE International*(2009-01-1859).
- [6] Cantore, G., Paltrinieri, F., Tosetti, F., and Milani, M., 2008. “Lumped parameters numerical simulation of a variable displacement vane pump for high speed ice lubrication”. *SAE International*(2008-01-2445).

- [7] Mancò, S., Nervegna, N., Rundo, M., and Armenio, G., 2004. "Modeling and simulation of variable displacement vane pumps for ic engine lubrication". *SAE International*(2004-01-1601).
- [8] Milani, M., Paltrinieri, F., Tosetti, F., Bianchini, A., Ferrara, G., and Ferrari, L., 2009. "Design and optimization of a variable displacemet vane pump for high performance ic engine lubrication, part 2". *SAE International*(2009-01-1064).
- [9] Bianchini, A., Ferrara, G., Ferrari, L., Paltrinieri, F., Tosetti, F., and Milani, M., 2009. "Design and optimization of a variable displacemet vane pump for high performance ic engine lubrication, part 1". *SAE International*(2009-01-1045).
- [10] Novi, N., Squarcini, R., and Frendo, F., 2009. "Dynamic and kinematic evaluation of automotive variable displacement vane pumps for reliability characterization". *SAE International*(2009-01-0191).
- [11] Fiebig, W., and Heisel, U., 1993. "Comportamento dinamico di pompe a palette a regolazione di pressione". *Oleodinamica-Pneumatica*, September, pp. 122–131.
- [12] Borghi, M., Milani, M., and Valentini, A., 1999. "Pompe a palette - sperimentazione e calcolo". *Oleodinamica Pneumatica Lubrificazione*, June, pp. 76–84.
- [13] Bussola, R., and Tiboni, M., 1999. Parametrical identification for a complex mathematical model of indexing cam mechanisms by means of genetic algorithm. Tech. rep., Eurogen, Jyväskylä, Finland, May 30 - June 3.

DESIGN AND CONTROL OF NEW ROBOTIC JOINTS FOR REHABILITATION EXOSKELETONS

Rocco Vertechy, Antonio Frisoli, Massimiliano Solazzi, Andrea Dettori,
Massimo Bergamasco
PERCRO Laboratory,
Scuola Superiore Sant'Anna, Italy
E-mail: {r.vertechy, a.frisoli, m.solazzi, a.dettori, m.bergamasco}@sssup.it

Abstract. *A novel upper-limb exoskeleton has been developed at PERCRO laboratory for the robotic rehabilitation of stroke patients. The exoskeleton architecture is modular and based on custom-made flexible robotic joints, each comprising a geared transmission with large reduction ratio and an integrated joint-torque sensor. The robotic joints are provided with an original Linear-Quadratic-Gaussian regulator that makes it possible to overcome the torque control issues associated to both joint flexibility and measurement noise. In this paper, the architecture and the joint-torque controller of the developed exoskeleton are presented along with some experimental results that validate the adopted design and control solutions.*

Keywords: *exoskeletons, robotic rehabilitation, flexible joints, torque control.*

1. INTRODUCTION

The reduction in motor functionality of upper extremities is one of the most common impairments resulting from spinal cord injuries, occupational and sport injuries, strokes and other disease processes. In these cases, partial or full sensorimotor recovery can be achieved through intense and constant rehabilitation protocols based on active and highly repetitive exercises [1-3]. Traditional treatment methods rely on laborious manual procedures carried out by highly skilled therapists. In the last decade, Robotic Assisted (RA) and Virtual Reality Based (VRB) rehabilitation has been proposed [4] primarily for improving treatment outcomes, but also for diminishing therapists' workload and overall healthcare expenses (for both patient and healthcare providers).

Tangible features and benefits RA and VRB rehabilitation may provide are:

- Increased repeatability in the execution of the rehabilitative exercise;
- Delivered care less dependent on the skill, attention and fatigue level of the therapist;
- Real-time measurement and visualization of patient physiological data and performance indices;
- Immediate assessment of the rehabilitation progress by direct comparison of actually measured and previously stored physiological data and performance indices;
- Provision of a more stimulating rehabilitation environment for the patient;
- Increased duration and frequency of the therapy sessions which are likely to improve therapy efficacy;

- Possible execution of parallel therapy sessions under the surveillance of a single therapist (eventually located in a remote site).

So far, pilot studies have already demonstrated that RA and VRB rehabilitation is beneficial for the recovery of upper limb function [5-9], and a number of robotic rehabilitation devices have been developed that are either Cartesian or exoskeletal (extensive surveys of such devices can be found in [9-13]).

As compared to Cartesian devices, rehabilitation exoskeletons make it possible:

- a) To apply controlled forces at each joint or link of the patient upper extremity (e.g. arm, forearm, shoulder, elbow and wrist) and to modulate their magnitude depending on patient deficiency and exercise type;
- b) To fully or partially compensate for the weight and the inertia of the patient upper extremity;
- c) To replicate most (roughly 70%) of the patient upper extremity healthy workspace;
- d) To prevent motor compensatory strategies frequently employed by impaired patients for circumventing their motor deficiencies;
- e) To be worn rather quickly;
- f) To be fitted (relatively easily) to rather different patient limb sizes by adjusting the exoskeleton link lengths.

On the basis of the experience gained at PERCRO laboratory during the clinical evaluation [9] of the L-EXOS [14], an exoskeleton prototype with flexible robotic joints, hereafter called RehabExos, has been conceived for the rehabilitation of upper extremities. In this paper, the architecture and the robotic-joint specifications of the recently developed RehabExos are described together with a method for the torque-control of its flexible robotic joints.

2. THE REHABHEXOS

The user requirements for the RehabExos have been defined during the clinical evaluation of the L-EXOS [9]. Besides the specifications indicated at points (a)-(f) of Section 1, the RehabExos features:

- Exercisable force at the end-effector of 100 N, in every direction and in each point of the workspace;
- Redundant sensors for patient safety;
- Modular design for maintenance ease;
- Limited back-drivability of the system at motor power-off for safety and good transparency at motor power-on;
- Possibility of easily changing configuration for the right and left arm;
- Serial kinematics with 5 Degrees of Freedom (DoFs), the first 4 DoFs being actuated;
- Position sensor on each of the five DoFs;
- Isomorphism with the human arm kinematics.

In the following sub-sections, the kinematic architecture and the robotic-joint specifications of the RehabExos are described in more details.

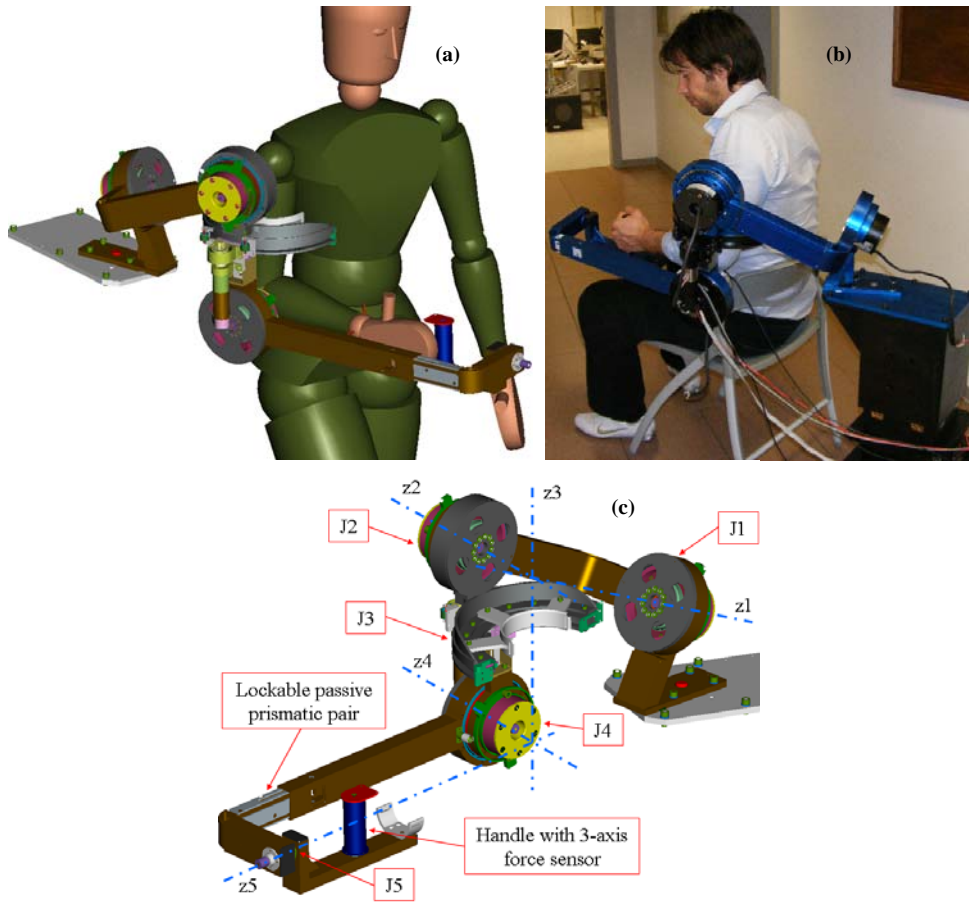


Figure 1. RehabExos: (a) CAD model worn by a mannequin; (b) Manufactured prototype worn by a human user; (c) CAD model showing the kinematic architecture.

KINEMATIC ARCHITECTURE

Figures (1.a) and (1.b) show the CAD model and the manufactured prototype of the RehabExos worn by a mannequin and a human user respectively. In this system, the patient can be constrained to the exoskeleton links at the middle of the arm and at the level of the forearm. In addition, the patient can grab the robot end-effector, which is a handle with a 3-axis force sensor.

Due to the effectiveness demonstrated by the L-EXOS architecture in performing rehabilitative tasks [9], the same kinematics has been replicated on the RehabExos. That is, as depicted in Fig. (1.c), the RehabExos has a shoulder joint which is fixed in space and composed by three actuated joints $J1$, $J2$ and $J3$, an actuated elbow joint $J4$ and a passive revolute joint $J5$ allowing for wrist pron/supin-ation. Referring to Fig. (1.c), the center of the spherical motion of the shoulder coincides with the point of intersection of axes $z1$, $z2$ and $z3$, while the elbow rotation axis is identified with $z4$. The intersection between $z3$ and $z4$ locates the elbow position.

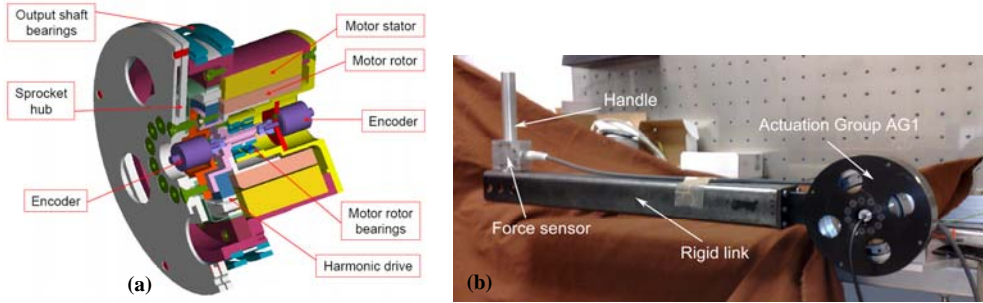


Figure 2. Actuation Group 1 (AG1): (a) CAD section exploded view with indication of the principal elements; (b) Manufactured prototype in its test-rig.

Different from the L-EXOS, the RehabExos forearm is composed by two links serially connected by a lockable passive prismatic pair. This design provides the RehabExos with some degree of adaptation to different sizes of patient limbs.

Consistent with the majority of the developed exoskeletons and with the successful rehabilitation protocols performed using the L-EXOS [9], also the RehabExos does not explicitly compensate for the translations of the patient gleno-humeral joint, which are possible in humans due to the presence of the closed-loop kinematic chain featuring the sternum, the clavicle, the scapula and the thorax links, and the sterno-clavicular, the acromio-clavicular and the scapulo-thoracic joints. In fact, owing to the limits of available actuator technologies, accounting for gleno-humeral joint translations is not worth at present. Note indeed that the potential increase in therapy benefits does not parallel the consistent augment in costs, weight and control complexity, which are required to build exoskeletons featuring such motion compensations.

Because of the serial kinematic architecture, both the torque and the dynamic performance which are required by the different joints of the exoskeleton are not the same. Even so, the RehabExos implements the same actuation group (hereafter called AG1) for the three joints J1, J2 and J4, and uses a different solution (hereafter called AG2) for J3. The choice of pursuing modularity has been made in order to reduce the mechanical complexity and the costs of the exoskeleton prototype. In the following subsections a detailed analysis of the robotic joints AG1 and AG2 is provided.

Design of Actuation Group 1

As already described, the three degrees of freedom provided by joints J1, J2 and J4 are motorized through identical robotic joints of type AG1. In order to comply with speed, torque and safety requirements of typical rehabilitative applications, the AG1 design specifications have been the following:

- 10rpm maximum output velocity;
- 120Nm nominal maximum output torque for actively sustaining exoskeleton links and patient limbs, and for generating the desired contact forces;
- Limited joint back-drivability at motor power-off for guaranteeing patient safety;
- Joint torque sensorization to derive forces at the exoskeleton-patient contact points, as well as torques in the patient joints;
- Redundant sensorization to monitor system integrity;

- Limited mechanical complexity to ease maintenance and to reduce costs.

The final design of AG1 is depicted in Fig. (2a). It consists of a custom-made frameless brushless torque motor integrating a Harmonic Drive (HD) component set, two redundant optical encoders, two thin output shaft bearings, and a torque sensor consisting of two fully balanced strain gauge bridges placed on different beams of a thin planar sprocket hub. The sprocket hub is connected to the HD output at one side and to the AG1 output shaft at the other. This design makes the torque sensor insensitive to non-torsional components of the external loads acting on the output shaft. It has to be mentioned that joint torque could have been measured by placing the strain gauges directly on the HD flexspline [15]. We preferred the addition of a sprocket hub since the chosen HD is very compact, thus it is very difficult to be instrumented, and since HD-based torque sensing inherently introduces significant measurement errors which require a large number of strain gauges and appropriate filtering techniques to be compensated [16, 17]. The sprocket hub is made of steel (AISI 630), while the motor housing and the non-commercial components are made of aluminium alloy (Ergal 7075 T6).

As for the component specifications, the motor has a maximum velocity of 1000rpm and a nominal stall torque of 2Nm; the HD has a very short axial length (50% of standard HD), reduction ratio of 100 and a limiting average output-torque of 151Nm; the encoders have 5000 pulses/revolution; the torque sensor has a measurement range of $\pm 120\text{Nm}$ (sprocket hub deformation is limited by mechanical stops); the overall joint stiffness (i.e. the in-series stiffness of HD and sprocket hub) is 6.3kNm/rad. The overall weight of AG1 is 3.7Kg.

The manufactured AG1 prototype embedded in a test-rig is shown in Fig. (2.b). The test-rig consists of a fixed frame attached to the motor stator, a rigid link attached to the AG1 output shaft, and a handle. The rigid link inertia is chosen so as to simulate the exoskeleton inertia in the most adverse configuration. The handle has a single axis force sensor, which enables the measurement of the active forces (i.e. those producing mechanical work) a user can exchange with the link.

Design of Actuation Group 2

The exoskeleton joint J3 consists of a different actuation group (AG2). The subsystem design specifications were:

- Remote exoskeleton axis of rotation to mimic the int/ext-ernal rotary motion of the patient shoulder and to provide the patient arm with a controlled torque along the same axis;
- Easy accommodation of the patient arm;
- Limited joint back-drivability at motor power-off;
- Joint torque sensorization;
- 10rpm maximum output velocity;
- 80Nm nominal maximum output torque;
- Limited mechanical complexity to ease maintenance and reduce costs.

The final design of AG2 is depicted in Fig. (3.a). It consists of a custom-designed open circular guide rail and a roller slider (using eight commercial rollers), a commercial geared DC motor integrating an optical encoder, a cable drive transmission and a force sensor consisting of a balanced strain gauge bridge placed on a parallelogram element. The motor

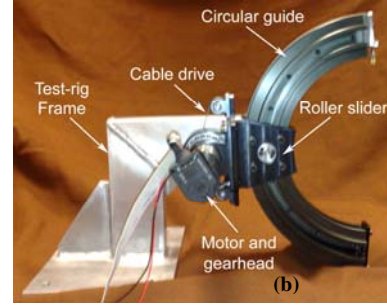
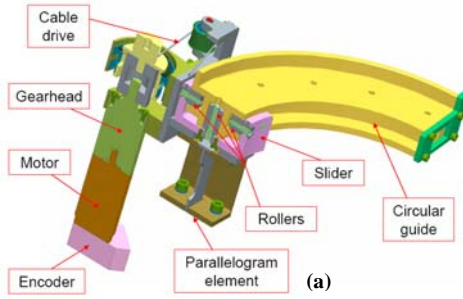


Figure 3. Actuation Group 2 (AG2): (a) CAD section exploded view with indication of the principal elements; (b) Manufactured prototype in its test-rig.

stator is fixed to the roller slider, while the motor gear head output is connected to the circular guide via the cable drive. The parallelogram element of the force sensor is made of steel (AISI 630), while the non-commercial components are made of aluminium alloy (Ergal 7075 T6). The guide rails are hard-anodized to provide wear resistance.

As for the component specifications, the motor has a maximum velocity of 8000rpm and a nominal stall torque of 0.105Nm; the gearhead reduction ratio is 33, the cable drive reduction ratio is 23; the force sensor makes it possible to measure a joint torque within the range $\pm 80\text{Nm}$ (parallelogram element deformation is limited by a mechanical stop). The overall weight of AG2 is 2.2Kg.

The manufactured AG2 prototype embedded in a test-rig is shown in Fig. (3.b).

Experimental Validation and Characterization of the Actuation Groups

The robotic joint prototypes AG1 and AG2 have been tested in properly predisposed test-rigs (shown in Figures (2.b) and (3.b) respectively). The purpose of the experimental tests was the evaluation of:

- The correct integration of the employed electrical and mechanical components;
- The energy efficiency of the employed mechanical transmissions;
- The overall friction losses;
- The effective inertia and gravity loads.

Exact knowledge of the aforementioned information is fundamental for the performance maximization of the control system of the actuation groups and of the overall exoskeleton.

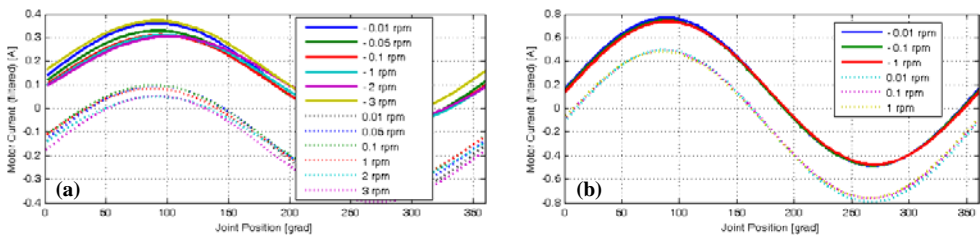


Figure 4. Motor current vs. output link position for different velocities of AG1: (a) with no external load; (b) with a 7.4Kg load applied to output link end.

Experimental results validated the design hypotheses. For the sake of conciseness, only the experimental data measured for AG1 are reported in the sequel. In particular, Figures (4.a) and (4.b) report, for different output link speeds, the motor currents required to maintain in uniform rotation the output link (whose length is 0.604m and weight is 4.15Kg) with no external load and with a 7.4Kg load applied at the end of the output link respectively. These data show that:

- The overall mechanical energy efficiency of AG1 varies between a minimum of 0.75 (for no external load) to a maximum of 0.85 (for the maximum external load);
- The friction losses of AG1 are nearly constant and independent of output link position and velocity and of external load magnitude. They can be compensated by a constant motor current with sign equal to that of the motor rotor speed.

3. JOINT-TORQUE CONTROL OF REHABEXOS ACTUATION GROUPS

As previously described, to enhance the interaction performance between RehaExos and its human wearer (i.e. superior force/torque and impedance modulation), the custom made robotic joints AG1 and AG2 have been equipped with integrated Joint-Torque Sensors (JTS). The availability of these internal sensors indeed enables:

- Closing a stable high-bandwidth torque inner loop around the joint which is scarcely affected by the robot link variable inertia [18, 19];
- Suppressing the robot vibrations produced by the inherent transmission compliance [20-22];
- Reducing internal disturbance torques caused by actuator and reducer (for instance friction losses, actuator torque ripples and gear teeth wedging actions [23]);
- Measuring externally applied forces/moments and complex nonlinear dynamic interactions between joint and other robot links [24].

However, JTS increases robot compliance (i.e. reduces its natural frequency) which asks for advanced control techniques to be compensated. To cope with this, we have developed a Linear-Quadratic-Gaussian (LQG) Joint-Torque-Feedback (JTF) regulator which is based on a state-space lumped-parameter model of flexible actuation groups. In the following sub-sections, this LQG-JTF regulator is first described. Then, experimental results acquired on the AG1 prototype are presented to show the good performances of this LQG-JTF regulator in the torque control of flexible robotic joints.

Model for Flexible Robotic Joints

The design of a torque controller for a flexible robotic joint (alike those shown in Figures 2 and 3) requires an adequate model of the system dynamics. A possible lumped parameter model is given in Fig. (5.a).

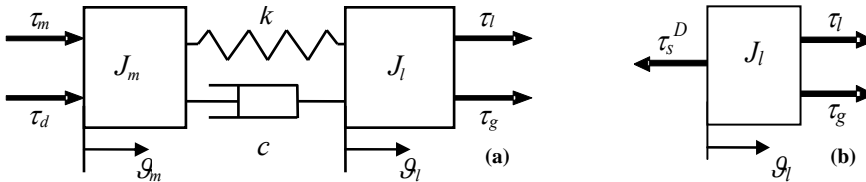


Figure 5. Flexible Robotic Joint: (a) Lumped parameter model; (b) Lumped parameter model of the flexible joint with JTF control.

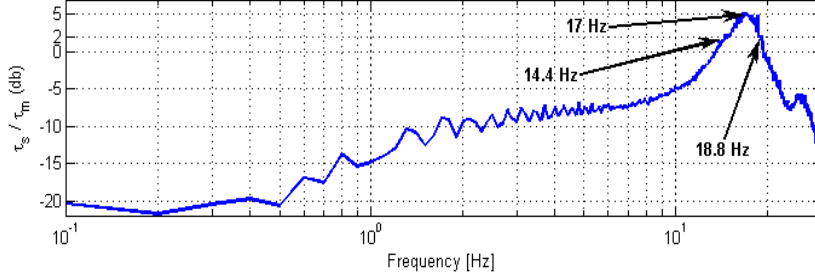


Figure 6. Experimental open-loop response of AG1: joint sensor torque vs. motor torque command.

In the model, J_m and J_l are the overall moments of inertia of the elements which are fixed with respect to motor rotor and joint output link, ϑ_m and ϑ_l are the angular positions of motor rotor and joint output link, τ_m and τ_d are the internal commanded and disturbance (which accounts for friction and ripple effects of both motor and harmonic drive) torques acting directly on the motor rotor, τ_g and τ_l are the external torques acting directly on the joint output link which are due to the gravitational field and to the contact interaction with the environment (for instance with a human user), k and c are the overall stiffness and damping coefficients of the flexible joint (principally due to the harmonic drive and the torque sensor). Owing to the presence of a mechanical transmission (i.e. the harmonic drive), note that all the aforementioned quantities are considered as being reduced to the same moving shaft, for instance the joint output link.

The dynamics of the system depicted in Fig. (5.a) is governed by the following equations

$$\begin{cases} J_m \ddot{\vartheta}_m = \tau_m + \tau_d + k(\vartheta_l - \vartheta_m) + c(\dot{\vartheta}_l - \dot{\vartheta}_m) \\ J_l \ddot{\vartheta}_l = \tau_l + \tau_g - k(\vartheta_l - \vartheta_m) - c(\dot{\vartheta}_l - \dot{\vartheta}_m) \end{cases} \quad (1)$$

Assuming that the flexible joint is equipped with an internal torque sensor capable of measuring the elastic torque τ_s (hereafter called sensor torque), which acts between motor rotor and joint output link, i.e. $\tau_s = k(\vartheta_l - \vartheta_m)$, and by considering its time derivatives, the sensor torque dynamics reads as

$$\ddot{\tau}_s + \frac{c}{J} \dot{\tau}_s + \frac{k}{J} \tau_s = \frac{k}{J_l} \tau_l + \frac{k}{J_l} \tau_g - \frac{k}{J_m} \tau_d - \frac{k}{J_m} \tau_m \quad (2)$$

where $J = [J_m J_l / (J_m + J_l)]$. With reference to the test-rig of AG1 depicted in Fig. (2.b), the spectrum of the joint sensor torque τ_s in response to the motor torque command τ_m (both reduced to the output shaft of AG1) is reported in Fig. (6). This Bode magnitude plot is obtained experimentally by commanding τ_m with a chirp function. As Eqn. (2) predicts, the test-rig possesses a marked natural frequency at $\sqrt{k/J}/2\pi \approx 17\text{Hz}$. From Fig. (6), use of the Half-Power Bandwidth method returns $c = 11.8\text{Nms/rad}$ as the overall damping coefficient of the flexible joint (this value has also been validated via the Logarithmic Decrement method).

Joint Torque Feedback Control for Flexible Robotic Joints

Environment interacting robots, in particular haptic and rehabilitation devices, are machines which modulate the forces/torques they exchange with the surroundings, in particular the human user. Referring to the one-degree-of-freedom flexible joint schematized in Fig. (5.a), this reduces to the regulation of τ_l by the action on τ_m . Equation (2) highlights that while the dynamics of the sensor torque τ_s can be shaped as desired by designing a proper feedback law for the commanded motor torque τ_m , conversely the external torque τ_l is not controllable via τ_m . Besides, the availability of a stable, fast and accurate torque control loop on τ_s (i.e. a JTF control) can enhance enormously the adjustment of the external torque τ_l . Indeed, by forcing $\tau_s \approx \tau_s^D$ and $\dot{\tau}_s \approx 0$, where τ_s^D is a desired sufficiently smooth-varying sensor torque, the second equation of Eqn. (1) tends to $J_l \ddot{\theta}_l \approx \tau_l + \tau_g - \tau_s^D$, which reduces the dynamics of the flexible joint to that of a direct drive system (alike the model described in Fig. (5.b)) having the desired torque τ_s^D as command. Then, the design of τ_s^D as a well-behaved function of θ_l (along with its time derivatives) and of τ_g (which can be easily calculated as a function of θ_l) enables the regulation of the interaction torque τ_l as well as the contact impedance of the joint output link. For instance, if the angular acceleration of the output link is available (via either direct measurement or an estimator), the adoption of a stable fast and accurate JTF control with $\tau_s^D \approx (\tau_l^D + \tau_g - J_l \ddot{\theta}_l)$ can be used to force $\tau_l \approx \tau_l^D$, where τ_l^D is the desired joint output link torque. In addition, whenever a direct measurement of τ_l is available, a similar result (i.e. $\tau_l \approx \tau_l^D$) can be obtained by setting the JTF control with $\tau_s^D = [\tau_l^D + \tau_g + k_d (\tau_l^D - \tau_l)]$, where k_d is an appropriate feedback error gain of an outer torque loop. Note that for most haptic applications, and in particular for rehabilitation, which usually involve motions with rather limited accelerations, the simple choice $\tau_s^D = (\tau_l^D + \tau_g)$ suffices.

State-Space Controller for Flexible Robotic Joints

In this section, the design of a stable JTF controller that is capable of tracking $\tau_s \approx \tau_s^D$ and $\dot{\tau}_s \approx 0$ with great accuracy and responsiveness is performed via a state-space approach [25]. For the purpose of JTF control, Eqn. (2) is recast in the following state-space system

$$\dot{\mathbf{e}} = \mathbf{A}\mathbf{e} + \mathbf{B}\tau_m + \mathbf{E}\tau_0 \quad (3.1)$$

where τ_m is the control input,

$$\mathbf{e} = \begin{bmatrix} \tau_s - \tau_s^D \\ \dot{\tau}_s \end{bmatrix}, \quad \tau_0 = \begin{bmatrix} \tau_{lgd} \\ \tau_s^D \end{bmatrix}, \quad \dot{\tau}_0 = \mathbf{0} \quad (3.2)$$

are the controllable and the uncontrollable (i.e. exogenous input) variables, $\tau_{lgd} = [\tau_l + \tau_g - \tau_d (J_l / J_m)]$, and

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -k/J & -c/J \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ -k/J_m \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 0 & 0 \\ k/J_l & -k/J \end{bmatrix} \quad (3.3)$$

Note that Eqn. (3) assumes $\dot{\tau}_s^D \approx 0$ (i.e. τ_s^D is a sufficiently smooth-varying desired sensor torque). For more demanding applications, whenever $\dot{\tau}_s^D$ is available, the controllable variable $\dot{e} = \dot{\tau}_s$ can be replaced by $\dot{e} = \dot{\tau}_s - \dot{\tau}_s^D$ whilst \mathbf{E} and τ_0 augmented to account $\dot{\tau}_s^D$ as supplementary exogenous input variable. According to Eqn. (3), the following full-state feedback law is obtained

$$\tau_m = \tau_{m,1} = -\left[k_1 \left(\hat{\tau}_s - \tau_s^D \right) + k_2 \hat{\tau}_s + k_3 \hat{\tau}_{lgd} + k_4 \tau_s^D \right] \quad (4)$$

where k_1 , k_2 , k_3 and k_4 are the controller gains, while $\hat{\tau}_s$, $\hat{\tau}_s^D$, $\hat{\tau}_{lgd}$ and τ_s^D are the estimates of both controllable and exogenous state-space variables. Note that whenever a reliable model $\hat{\tau}_d$ is readily available for the motor disturbance τ_d , Eqn. (4) can be replaced by

$$\tau_m = \tau_{m,2} = -\left[k_1 \left(\hat{\tau}_s - \tau_s^D \right) + k_2 \hat{\tau}_s + k_3 \hat{\tau}_{lg} + k_4 \tau_s^D \right] - \hat{\tau}_d \quad (5)$$

where $\hat{\tau}_{lg}$ is the estimate of the overall external torque $\tau_{lg} = (\tau_l + \tau_g)$, which acts on the joint output link only.

In principle, the appropriate choice of the controller gains makes it possible to shape at will the dynamic response of τ_s to τ_s^D . In order to mediate between the exigencies of fast and accurate system response and the saturation limits of the control command τ_m , the gains k_1 , k_2 , k_3 and k_4 are selected in the framework of Linear-Quadratic (LQ) optimum control [25]. In particular, k_1 , k_2 , k_3 and k_4 are chosen as the steady-state gains that minimize the following performance integral

$$LQ = \int_t^T \left[q_1^2 \left(\tau_s - \tau_s^D \right)^2 + q_2^2 \dot{\tau}_s^2 + \tau_m^2 \right] \cdot dt \quad (6)$$

where q_1 and q_2 are penalty parameters relative to the cost of control. As for their physical significance, the choice of q_1 influences directly the system tracking error e (i.e. larger q_1 leads to smaller e), while the choice of the ratio q_2/q_1 influences directly system stability and responsiveness (i.e. rising q_2/q_1 increases stability and reduces time-response). Augmenting q_1 and q_2 enlarges controller effort. The practical mathematical procedure for calculating k_1 , k_2 , k_3 and k_4 for a given system (i.e. equations (4) and (6)) and chosen penalty parameters is detailed in [25].

As compared to the most popular JTF laws, the controller described by either Eqn. (4) or Eqn. (5) (hereafter called JTF1) differs from that described in [26, 27] (hereafter called JTF2) since there $k_3 = 0$ and $k_4 = (1 + J_m/J_l)$, and from that described in [28, 29] (hereafter called JTF3) since there $k_3 = 0$ and $k_4 = 1$. That is, the controller proposed here can be considered as a generalization of those given in [26-29].

Optimum Observer for Flexible Robotic Joints

Beside the availability of τ_s^D , the JTF control scheme proposed in the previous section requires the knowledge of the state-space variable estimates $\hat{\tau}_s$, $\hat{\tau}_s^D$ and $\hat{\tau}_{lgd}$ (or $\hat{\tau}_{lg}$ and $\hat{\tau}_d$ if Eqn. (5) is used instead of Eqn. (4)). Here, a Kalman filter (optimum observer [25]) is

proposed to infer $\hat{\tau}_s$, $\hat{\tau}_s^m$ and $\hat{\tau}_{lgd}$ from a single noisy measurement τ_s^m , $\tau_s^m = (\tau_s + w)$, where w is the noise affecting the torque sensor embedded in the flexible joint (the major contribution to w is usually the electrical noise excited by the motor driver switching electronics).

For the purpose of variable estimation (both controllable states and exogenous input), the dynamics governed by Eqn. (2) and Eqn. (4) can be described by the following meta-system

$$\begin{cases} \dot{\boldsymbol{\tau}} = \mathbf{A}'\boldsymbol{\tau} + \mathbf{B}'\tau_{m,1} + \mathbf{F}v \\ y = \tau_s^m = \mathbf{C}\boldsymbol{\tau} + w \end{cases} \quad (7.1)$$

where $\boldsymbol{\tau}^T = [\tau_s \quad \dot{\tau}_s \quad \tau_{lgd}]$ is the meta-state vector, w and v are white noise processes with known variances W and V , $\mathbf{C} = [1 \ 0 \ 0]$ and

$$\mathbf{A}' = \begin{bmatrix} 0 & 1 & 0 \\ -k/J & -c/J & k/J_l \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{B}' = \begin{bmatrix} 0 \\ -k/J_m \\ 0 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (7.2)$$

Note that in order for the Kalman filter to estimate the exogenous input τ_{lgd} , the dynamics of such an unknown variable is assumed as the Wiener process (i.e. a non-stationary random process) $\dot{\tau}_{lgd} = v$, which, in practice, only forces τ_{lgd} to be continuous and varying with independent increments. According to Eqn. (7), the following optimum observer is obtained

$$\hat{\boldsymbol{\tau}} = \mathbf{A}'\hat{\boldsymbol{\tau}} + \mathbf{B}'\tau_{m,1} + \mathbf{L}(y - \mathbf{C}\hat{\boldsymbol{\tau}}) \quad (8)$$

where $\hat{\boldsymbol{\tau}}^T = [\hat{\tau}_s \quad \hat{\dot{\tau}}_s \quad \hat{\tau}_{lgd}]$ is the meta-state vector estimate and \mathbf{L} is the steady-state Kalman gain vector. The practical mathematical procedure for calculating \mathbf{L} for a given system (i.e. once Eqn. (7) is known) and noise process variances (i.e. W and V) is detailed in [25]. In this regard, note that the variance W is evaluated directly from the torque sensor measure, while the ratio V/W is chosen so as to compromise between estimator accuracy and precision (i.e. larger V/W leads to more accurate but less precise estimates).

Note that when a reliable model $\hat{\tau}_d$ is readily available for the motor disturbance τ_d (such that Eqn. (5) is used instead of Eqn. (4) for JTF control), the substitution in Eqn. (8) of $\tau_{m,1}$ with $(\tau_{m,2} + \hat{\tau}_d)$ enables to find $\hat{\tau}_{lg}$ instead of $\hat{\tau}_{lgd}$.

It is worth to mention that the JTF implementations, which can be found in the literature for the torque control of flexible joints, do not consider the possibility of reconstructing the exogenous variable τ_{lgd} (or τ_{lg} if Eqn. (5) is used instead of Eqn. (4)) from the measurement of the built-in joint-torque sensor. Indeed, just minimum order observers for the estimation of τ_d have only been devised and employed which require the additional measurement of either motor position θ_m [27, 30, 31, 32] or motor velocity $\dot{\theta}_m$ [22]. Moreover, in these JTF implementations, the estimates $\hat{\tau}_s$ and $\hat{\dot{\tau}}_s$ are obtained by simply feeding the torque sensor measurement y through basic low-pass and derivative filters, instead of reconstructing them via the optimal observer given by Eqn. (8). This may

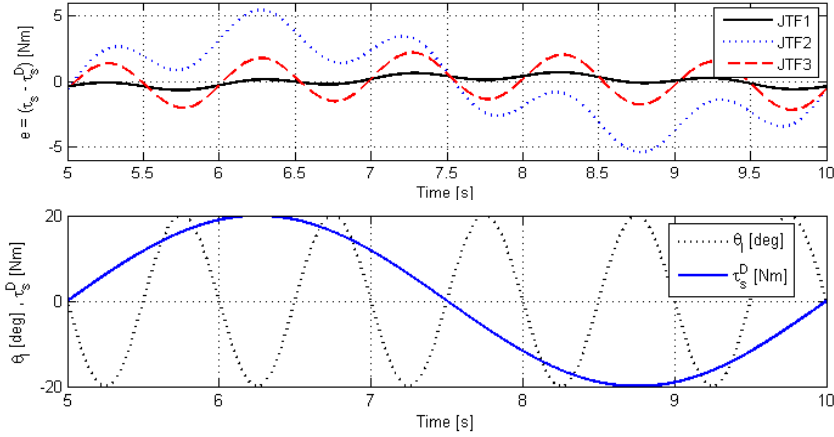


Figure 7. Comparison between different Joint Torque Feedback (JTF) laws (simulation results).

degrade JTF control performance since built-in joint-torque sensors are usually affected by noise w having significant variance.

Simulation and Experimental Results

The combination of Eqn. (4) (or Eqn. (5)) with Eqn. (8) gives a LQG torque regulator for flexible robotic joints. In this sub-section, such a regulator is employed for the control of the robotic joint AG1 depicted in Fig. (2.b). The properties of AG1 are (all quantities reduced to the joint output shaft): motor shaft inertia $J_m = 2.5\text{Kg}\cdot\text{m}^2$; rigid link inertia $J_l = 0.7\text{Kg}\cdot\text{m}^2$; overall joint torsional stiffness $k = 6.3\text{kNm/rad}$; overall joint torsional damping $c = 11.8\text{Nms/rad}$; torque sensor variance $W = 0.12\text{Nm}$ (measured with the motor driver turned on). According to these properties, the choice $V = 10^4$ yields the following Kalman gain vector $L = [199, 19854, 289]^T$, whereas the choice $q_1 = 500$ and $q_2 = 0.1$ gives the controller gains $k_1 = -18.3$, $k_2 = -0.33$, $k_3 = -2.81$ and $k_4 = 3.62$.

The synthesized controller is first tested in simulation to verify the performance of the proposed novel control law JTF1 with respect to the existing JTF2 [26, 27] and JTF3 [28,29]. In the simulation, the following working conditions are considered: $\tau_d = \tau_g = W = 0$, $\tau_s^D = [20 \sin(0.4\pi t)]\text{Nm}$ (t being the elapsed time), $\theta_l = [20 \sin(2\pi t)]\text{deg}$ (that is, the controller attempts to track τ_s^D while the output link is moved according to a given trajectory; measurement errors are not considered here). Simulation results are reported in Fig. (7), which highlights the better performance (i.e. smaller tracking error e , see upper plot) of JTF1 with respect to both JTF2 and JTF3. In particular the tracking error obtained with JTF1 is $|e| \leq 0.3\text{Nm}$.

Then, the proposed LQG regulator is assessed on the test-rig shown in Fig. (2.b). In the experiment, the joint axis is vertical so that $\tau_g = 0$. Moreover, disposing of the following experimentally determined analytical model for the motor friction torque $\hat{\tau}_d = [-13.7 \cdot \text{sign}(\dot{\theta}_m)(1 + \tanh(|\dot{\theta}_m|/2) - 4)/2]\text{Nm}$, the controller implements Eqn. (5) instead of Eqn. (4). Since $\tau_g = 0$, the Kalman filter makes it possible to get an estimate of the external torque τ_l (i.e. $\hat{\tau}_l$). In the experiment, the following desired torque is considered $\tau_s^D = [20 \sin(0.4\pi t)]\text{Nm}$, while the end link is moved in the range

$\theta_l \in [-60,60]\text{deg}$ by a human user who acts on the sensorized handle of the test-rig (see Fig. (2.b)). Experimental results are reported in Fig. (8), which highlight (third plot from the top) that the proposed controller is able to keep the torque sensor tracking error e within $\pm 0.5\text{Nm}$ (which is about the same value as that obtained in simulation and of the noise of the joint-torque sensor measure). The effectiveness of the Kalman filter in the estimation of τ_s and τ_l is shown, respectively, in the first two diagrams where the estimates $\hat{\tau}_s$ and $\hat{\tau}_l$ are plotted over the raw measurements τ_s^m and τ_l^m (τ_l^m being measured via the force sensor placed at the test-rig handle). The significant difference (roughly $\pm 3\text{Nm}$) between $\hat{\tau}_l$ and τ_l^m , which is also evidenced in the third plot from the top, is only due to the imperfection of the adopted friction model $\hat{\tau}_d$ (in practice, $\hat{\tau}_l$ depends on τ_l^m but also on the error between the real and the modeled motor disturbance torque).

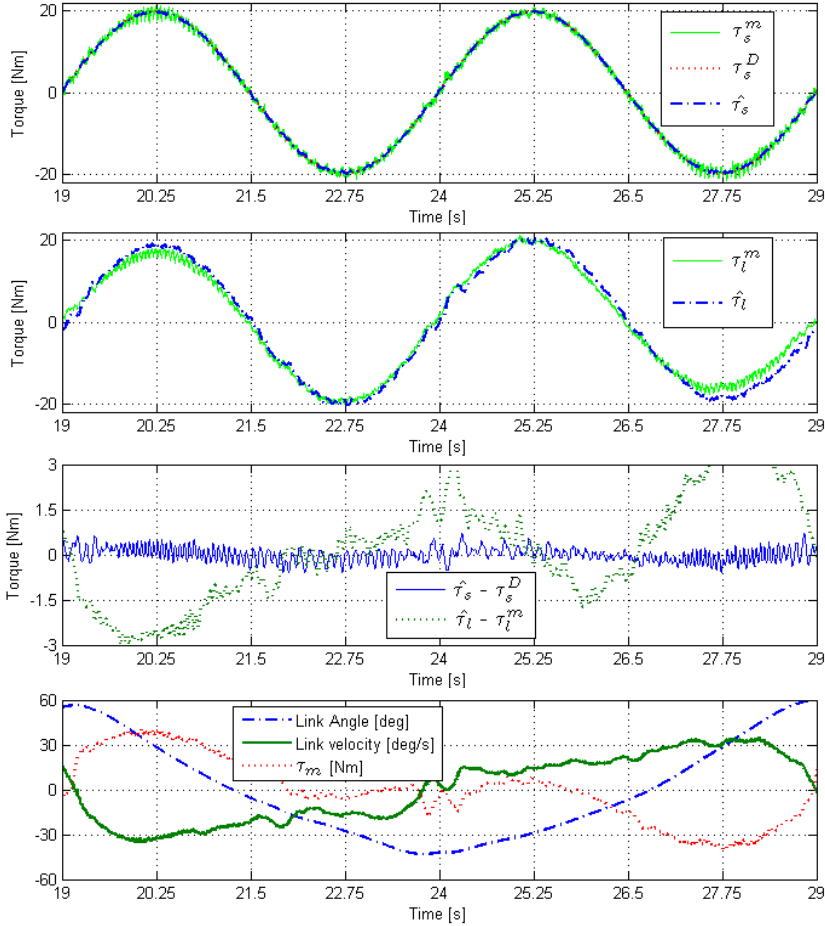


Figure 8. Performance of the proposed Linear-Quadratic-Gaussian (LQG) torque regulator (experimental results)

The bottom graph plots the position and velocity of the output link together with the commanded motor torque τ_m . As shown, τ_m is rather smooth and remains within $\pm 40\text{Nm}$ and does not exceed too much the maximum of the sum of desired sensor torque and motor friction torque (i.e. $\max(\tau_s^D + \hat{\tau}_d) \approx 33.7\text{ Nm}$).

4. CONCLUSIONS

A novel exoskeleton design has been presented for the rehabilitation of upper extremities. The exoskeleton is aimed at generating controlled contact forces not only at the exoskeleton handle, but also at its intermediate links. The novel exoskeleton design is modular and based on two different flexible robotic joints with integrated torque sensors, which have been conceived to provide the exoskeleton with scarce back-drivability at system power-off and accurate contact-force control when the system is on. To circumvent the control issues that are inherent in flexible robotic joints, a novel Linear-Quadratic-Gaussian regulator has been devised and implemented on each of the actuation groups the exoskeleton is composed by. As compared to other available joint-torque-feedback controllers, the proposed regulator makes it possible to overcome most of the issues that are related to the presence of noise in the joint-torque sensor measure and to the lack of knowledge of both the external and the internal torques acting on the robotic joint.

ACKNOWLEDGEMENTS

This work has been financially supported by the Italian Fondazione Monte Paschi.

REFERENCES

- [1] Butefisch, C., Hummelsheim, H., Denzler, P. and Mauritz, K.H., 1995. "Repetitive training of isolated movements improves the outcome of motor rehabilitation of the centrally paretic hand". *J. Neurol. Sci.*, **130**(1), pp. 59–68.
- [2] Diller, L., 2000. "Post-stroke rehabilitation practice guidelines". *International handbook of neuropsychological rehabilitation. Critical issues in neurorehabilitation*. New York: Plenum, pp. 167–182.
- [3] Barreca, S., Wolf, S.L., Fasoli, S., and Bohannon, R., 2003. "Treatment Interventions for the Paretic Upper Limb of Stroke Survivors: A Critical Review". *Neurorehabilitation and Neural Repair*, **17**(4), pp. 220–226.
- [4] Krebs, H.I., Hogan, N., Aisen, M.L. and Volpe, B.T., 1998. "Robot-aided neurorehabilitation". *IEEE Transactions on Rehabilitation Engineering*, **6**(1), pp. 75–87.
- [5] Reinkensmeyer, D.J., Kahn, L.E., Averbuch, M., McKenna-Cole, A., Schmit, B.D., and Rymer, W.Z. 2000. "Understanding and treating arm movement impairment after chronic brain injury: Progress with Arm Guide". *Journal of Rehabilitation Research and Development*, **37**(6).
- [6] Jack, D., Boian, R., Merians, A., Tremaine, M., Burdea, G., Adamovich, S., Recce, M. and Poizner, H. 2001. "Virtual reality-enhanced stroke rehabilitation". *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **9**(3), pp. 308–318.
- [7] Cardoso, L., daCosta, R., Piovesana, A., Costa, M., Penna, L., Crispin, A., Carvalho, J., Ferreira, H., Lopes, M. and Brandao, G., 2006. "Using Virtual Environments for Stroke Rehabilitation". *International Workshop on Virtual Rehabilitation*.
- [8] Stewart, J., Yeh, S., Jung, Y., Yoon, H., Whitford, M., Chen, S., Li, L., McLaughlin, M., Rizzo, A. and Winstein, C., 2006. "Pilot Trial Results from A Virtual Reality System Designed to

- Enhance Recovery of Skilled Arm and Hand Movements after Stroke”. *International Workshop on Virtual Rehabilitation*.
- [9] Frisoli, A., Montagner, L., Borelli, L., Salsedo, F. and Bergamasco, M., 2008. “A force-feedback exoskeleton for upper limb rehabilitation in Virtual Reality”. *Applied Bionics and Biomechanics*, **0**(0), pp.1-17.
 - [10] Tsagarakis, N. and Caldwell, D., 2003. “Development and Control of a 'Soft-Actuated' Exoskeleton for Use in Physiotherapy and Training”. *Autonomous Robots*, **15**(1), pp. 21-33.
 - [11] Nef, T. and Riener, R., 2005. “ARMin-Design of a Novel Arm Rehabilitation Robot”. *ICORR 2005, 9th International Conference on Rehabilitation Robotics*.
 - [12] Perry, J., Rosen, J. and Burns, S., 2007. “Upper-Limb Powered Exoskeleton Design”. *IEEE/ASME Transactions on Mechatronics*, **12**(4), pp. 408-417.
 - [13] Carignan, C., Naylor, M. and Roderick, S., 2008. “Controlling shoulder impedance in a rehabilitation arm exoskeleton”. *ICRA 2008, IEEE International Conference on Robotics and Automation*.
 - [14] Salsedo, F., Dettori, A., Frisoli, A., Rocchi, F., Bergamasco, M. and Franceschini, M., 2002. “Exoskeleton Interface Apparatus”.
 - [15] Hashimoto, M., 1989. “Robot motion control based on joint torque sensing”. *ICRA 1989, International Conference on Robotics and Automation*.
 - [16] Godler, I. and Hashimoto, M., 1998. “Torque control of harmonic drive gears with built-in sensing”. *IECON '98, Proceedings of the 24th Annual Conference of the IEEE Industrial Electronics Society*.
 - [17] Taghirad, H.D., Helmy, A. and Belanger, P.R., 1999. “Intelligent built-in torque sensor for harmonic drive system”. *IEEE Transaction on Instrumentation and Measurement*, **48**(6).
 - [18] Luh, J., Fisher, W. and Paul R., 1983. “Joint Torque Control by a Direct Feedback for Industrial Robots”. *IEEE Trans. Automan. Contr.*, **28**(2), pp. 153-161.
 - [19] An, C.H. and Hollerbach, J.M., 1983. “Dynamic Stability Issues in Force Control of Manipulators”. *Proc. IEEE Int. Conference Robotics and Automation*.
 - [20] Hashimoto, M., 1989. “Robot Motion Control Based on Joint Torque Sensing”. *Proc. IEEE International Conference Robotics and Automation*.
 - [21] Eppinger, S.D. and Seering, W.P., 1987. “Understanding Bandwidth Limitations in Robot Force Control”. *Proc. IEEE Int. Conf. Robotics and Automantion*.
 - [22] Zhang, G. and Furusho J., 1997. “Control of Robot Arms Using Joint Torque Sensors”. *Proc. of IEEE Int. Conf. on Robotics and Automation*.
 - [23] Pfeffer, L.E., Khatib, O. and Hake, J., 1989. “Joint Torque Sensory Feedback in the Control of the PUMA Manipulator”. *IEEE Trans. Robotics and Automation*, **5**(4), pp. 418-425.
 - [24] Kosuge, K., Takeuchi, H. and Furuta, K., 1988. “Motion Control of a Robot Arm Using Joint Torque Sensors”. *Proc. 27th Conf. Decision Control*.
 - [25] Friedland, B., 2005. *Control System Design: An Introduction to State Space Methods*. Dover Publications, New York.
 - [26] Hashimoto, M., Kiyosawa, Y. and Paul, R.P., 1993. “A Torque Sensing Technique for Robots with Harmonic Drives”. *IEEE Transactions on Robotics and Automation*, **9**(1), pp. 108-116.
 - [27] Hashimoto, M. and Kiyosawa, Y., 1998. “Experimental Study on Torque Control Using Harmonic Drive Built-in Torque Sensors”. *Journal of Robotic Systems*, **15**(8), pp. 435-445.
 - [28] Albu-Shaffer, A. and Hirzinger, G., 2000. “State Feedback controller for flexible joint robots: A globally stable approach implemented on DLR's light-weight robots”. *Proc. of 2000 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*.
 - [29] Ott, C., Albu-Shaffer, A., Kugi, A. and Hirzinger, G., 2008. “On the Passivity Based Impedance Control of Flexible Joint Robots”. *IEEE Transaction on Robotics*, **24**(2), pp. 416-429.
 - [30] Kaneko, K., Murakami, T., Ohnishi, K. and Komoriya, K., 1994. “Torque Control with Nonlinear Compensation for Harmonic Drive DC Motors”. *20th International Conference on Industrial Electronics, Control and Instrumentation*.

- [31] Kaneko, K., Suzuki, N., Ohnishi, K. and Tanie, K., 1991. "High Stiffness Torque Control for Geared DC Motor Based on Acceleration Controller". *Proc. of IEEE Int. Conf. on Industrial Electronics Instrumentation and Control*.
- [32] Kaneko, K., Komoriya, K., Ohnishi, K. and Tanie, K., 1992. "Accurate Torque Control for a Geared DC Motor based on Acceleration Controller". *Proc. of IEEE Int. Conf. on Industrial Electronics Instrumentation and Control*.

EXPERIMENTAL TESTS OF JOINT PROTOTYPES FOR UPPER LIMB POWERED PROSTHESES

Marco Troncossi

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: marco.troncossi@unibo.it*

Alessandro Rivola

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: alessandro.rivola@unibo.it*

Abstract. *This paper deals with the campaign of experimental tests that has been designed to characterize the prototypes of powered prosthetic articulations developed by the University of Bologna in collaboration with the INAIL Prosthesis Center. In particular the test bench, the test procedures and the software developed to investigate the efficiency and the vibro-acoustic behaviour of the mechanisms are presented.*

Keywords: *Upper limb prosthesis, Experimental characterization, Mechanical efficiency, Noise emission, Mechanical vibrations*

1. INTRODUCTION

The Department of Mechanical Engineering of the University of Bologna and the INAIL Prosthesis Center (Vigorso di Budrio, Bologna, Italy) have been collaborating for the development of new myoelectric upper limb prostheses addressed to patients with a high-level amputation. The research activity has produced a number of powered articulations so far, with the purpose of providing the patients having a high-level amputation with satisfactory solutions to significantly recover the functional autonomy in their everyday living (off-the-shelf components are indeed not adequate for the severest injuries [1]).

Such as for the development of any given technological product, the experimental testing of the prosthesis prototypes is necessary to characterize the actual performances of the devices with respect to the expected technical specifications and to assess their overall quality. Thus, a test rig and ad-hoc test procedures have been set up to deeply analyse our prosthesis prototypes with the final purpose to understand if critical parts exist that need to be improved and to get specific guidelines about the possible revision of the design (and the manufacturing process as well).

In particular, an experimental campaign focused on the objective assessment of the prosthetic articulations' velocity, payload, current draining, efficiency, and vibro-acoustic emissions has been designed. These aspects are very important for the prosthetic systems and a poor response could entail the rejection of the product by the final users (i.e. the amputees). The tests that we have planned for measuring these aspects make it possible to objectively assess the prosthetic prototypes' quality and to quantify the level of

improvement after possible modifications of the mechanisms (by comparing the values of the monitored parameters before and after the design modification). Moreover, the test procedures would allow the definition of some acceptability criteria in order to assess the suitability of products acquired from external providers with respect to our own specifications.

This paper illustrates the set up of the test rig and the architecture of the software that was implemented to retrieve the experimental data, and finally explains the use of the measured data. The focus is placed on two main algorithms of the data processing, developed for quantifying the mechanical efficiency of the mechanisms and to monitor the noise and vibrations emissions.

2. THE PROSTHESIS PROTOTYPES

The prosthetic arm prototype that can be assembled at the INAIL Prosthesis Center laboratories [Fig. 1(a)] is composed of: commercial myoelectric hand and wrist rotator (that are not of interest in this study), the INAIL elbow EMEI-20 (a quite mature “product” that has been already certified) [2], the prototype of an actuated shoulder articulation composed of two independent mechanisms [3], and –although not introduced in the artificial limb yet– the prototype of a powered humeral rotator [4] [Fig. 1(b, c)].

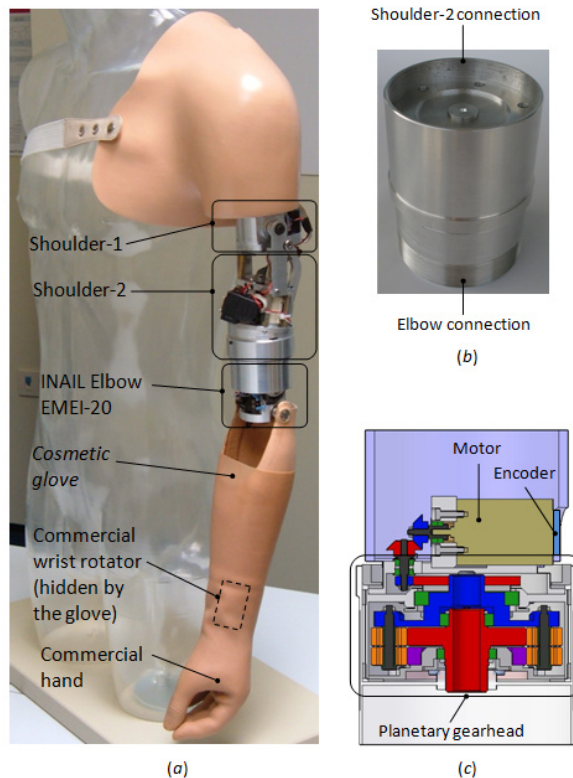


Figure 1. (a) Prosthesis prototype currently assembled at the INAIL Prosthesis Center laboratories. (b) Prototype of a new powered humeral rotator. (c) CAD model of the humeral rotator mechanism.

Table 1. Main characteristics of the DC motors that drive the prosthetic mechanisms.

Mechanism	Nominal Voltage V [Volt]	Maximum Power P ₂ [W]	No-load Speed n ₀ [rpm]	Stall Torque M _H [Nm]	Torque Constant k _M [Nm/A]	Friction Torque k _M [Nm]	Terminal Resistance R [Ohm]
Elbow	6	4.55	8200	0.021	0.00692	0.00020	1.94
Shoulder-1	6	4.55	8200	0.021	0.00692	0.00020	1.94
Shoulder-2	12	17.00	8100	0.080	0.01340	0.00100	1.90
Humeral Rotator	6	11.00	7100	0.059	0.00803	0.00028	0.81

Table 2. Main characteristics of the power transmission chains of the prosthetic prototypes.

Mechanism	Maximum Speed n _{follower} [deg/s]	Maximum Torque M _{follower} [Nm]	Reduction Ratio τ [-]
Elbow	40	6	1:984
Shoulder-1	35	2.5	1:1041
Shoulder-2	35	13	1:968*
Humeral Rotator	45	6.5	1:480

* The transmission chain includes an articulated linkage: its transmission ratio is variable with the mechanism configuration and 1:968 is the mean value.

The devices that actuate our upper limb prosthesis are electromechanical mechanisms powered by small DC electric motors. Due to the characteristics of these high-speed motors (Tab. 1), the power transmission chains of the devices must provide significant reduction ratios in order to have velocities and torques at the follower axes adequate for accomplishing the human manipulation tasks (Tab. 2). To this aim, many geared systems are used in the power transmission chains (ordinary stages, planetary gearheads, harmonic drives), which – despite their well known good qualities – can present problems, such as noise and/or low mechanical efficiency.

3. TEST PROCEDURES

Bench tests were studied to measure the actual performances (in terms of maximum payload and velocities, current draining...) of the single devices (tested one at a time) and to objectively quantify the prototypes quality. The definition of proper metrics to assess “quality” (physical quantities and/or indices defined ad-hoc) can be used to compare different design solutions, i.e. to assess the suitability/incidence of possible modifications made on a prototype by checking the values of the target parameters before and after the design changes.

In order to experimentally evaluate the overall structural quality of the prototype mechanisms, an energetic approach was chosen and the system efficiency was considered to be a good objective indicator to quantitatively assess the suitability of the mechanism design, the functional performance and the manufacturing accuracy. Two parameters were chosen, namely the mechanical efficiency of the power transmission chain (η_m) and the global efficiency of the whole electromechanical system (η_g), which takes into account also the contribution of the electric motor. Many details about this choice and the experimental determination of these parameters are reported in [5]. In the context of the present paper their operative definitions are sufficient in order to point out what should be measured for their quantitative evaluation:

$$\eta_g = \frac{L_u}{\int_{t_1}^{t_2} v(t) \cdot i(t) \cdot dt} \quad (1)$$

$$\eta_m = \frac{L_u}{\int_{t_1}^{t_2} \eta_{el}^* \cdot v(t) \cdot i(t) \cdot dt} \quad (2)$$

where $v(t)$ and $i(t)$ are respectively the instantaneous supply voltage [V] and the current [A] of the electric motor; L_u is the mechanical work outputted by the mechanism during the execution of a test trial (in the time interval $[t_1, t_2]$) and is theoretically computable once proper information are available [5]. The quantity η_{el}^* in the Eq. 2 is a sort of “instantaneous electric efficiency” of the DC that depends on the working conditions and can be computed as:

$$\eta_{el}^* = \frac{(k_M \cdot i(t) - M_R) \cdot \frac{\pi \cdot n(t)}{60}}{R \cdot i^2(t) + k_M \cdot i(t) \cdot \frac{\pi \cdot n(t)}{60}} \quad (3)$$

where $n(t)$ is the motor shaft velocity [rpm], k_M , M_R and R are the torque constant [Nm/A], the friction torque [Nm] and the terminal resistance [Ohm] of the motor respectively, whose nominal values are provided by the product datasheet (and reported in Tab. 1 as well). Therefore the only variables to be measured at the test bench for the determination of the global and mechanical efficiencies are the supply voltage $v(t)$, the current $i(t)$ and the motor shaft velocity $n(t)$ (note that the DC motors used in this application have a built-in incremental encoder).

The energy efficiency is not the only feature associated with the system quality: also the noise level is an appropriate tool of assessment, especially considering that a low noise is a fundamental requisite of prostheses, having the same importance – for instance – of payload and velocity. Therefore, we decided to quantify the noise emission of the prosthetic prototypes by acquiring the instantaneous air pressure-variation signals $[p(t)]$ with a microphone and calculating the Sound Pressure Level L_p (that is the simplest and most popular metrics used to define the noise level):

$$L_p = 20 \log_{10} \left(\frac{p_{rms}}{p_{ref}} \right) \text{ [dBA]} \quad (4)$$

where p_{ref} is the reference sound pressure (i.e. 20 μ Pa, considered as the threshold of human hearing) and p_{rms} is the root mean square value of the sound pressure being measured during one test trial. In order to make the pressure signals be consistent with the human perception of noise, the acquired pressure signals are filtered by means of the so-called A-Weighting Filter [6], and therefore the measurement unit associated with L_p is called “Decibel Adjusted” and indicated as dBA.

Since the noise emission is often associated with the mechanical vibrations of some mechanism parts, a proper investigation of this aspect could help in understanding how the designer can intervene to reduce the noise. Therefore also the vibrations affecting the mechanisms are measured, in terms of acceleration $a(t)$, by means of properly positioned transducers. The evaluation of both the efficiency and vibro-acoustic aspects is performed considering a simple reference test trial: the follower is driven from an initial position to a final different one

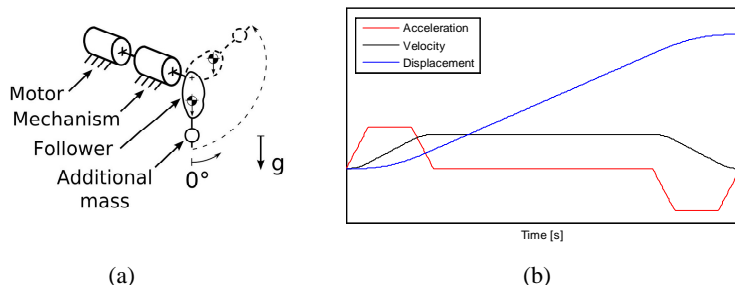


Figure 2. (a) Configuration of a prosthetic mechanism during a test trial, with the follower axis positioned in the horizontal direction, notwithstanding the real arrangement in the final artificial limb. (b) Qualitative trend of the motion law shapes.

(with initial and final velocity equal to zero) with a mono-directional flow of the energy, i.e. the mechanism actuator works only as a motor to equilibrate the external load which always has a resistant role. The external load is provided by the gravity force (chosen as the cheapest solution) acting on additional masses properly fixed to the follower [Fig. 2(a)]. Despite this simplicity, for the soundness of the retrieved results a certain accuracy in the control of the follower motion is needed in order to guarantee the accomplishment of properly designed trajectories [Fig. 2(b)] (note that the implementation of closed-loop control schemes is possible thanks to the encoders embedded in the motors). A number of tests with different payload and velocity are performed for each mechanism.

4. TEST BENCH

This paragraph describes the test rig that was designed and the corresponding software architecture implemented for the motion control and data acquisition.

Hardware

The experimental apparatus used for controlling the mechanism motion and acquiring the requested data is based on the National Instruments technology (Fig. 3). In particular, the following components are present:

- 1 Compact RIO NI cRIO 9022, with 8 slots CRIO-9114
- 1 Analog Input board NI 9234, with 4 channels (for the microphone and accelerometers)
- 2 Analog Input board NI 9201, with 4 channels (for the voltage and current signals)
- 1 Digital Input board NI 9425 board, with 32 channels (for the encoders' signals)
- 2 Digital Output board NI 9474, with 8 channels (for the generation of the PWM signals used for controlling the motors).

An intermediate board (Shoulder Board in Fig. 3) hosts electronic components needed for controlling the motor (i.e. the motor driver VNH2SP30 by STMicroelectronics) and measuring the current (i.e. current-sense amplifier MAX472 by Maxim Inc.). The voltage signal is detectable on a specific pin of the Shoulder Board that is directly connected to the analog input module NI 9201. Other accessories are present for the correct functioning of the system, e.g. for supplying energy or for transforming voltage to proper levels (by means

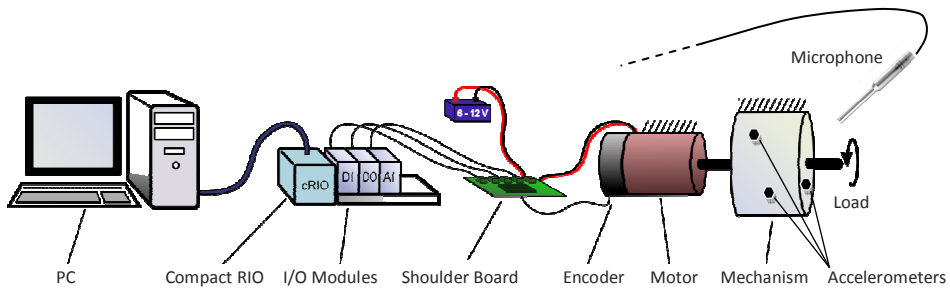


Figure 3. Layout of the experimental apparatus.

of photocouplers). The sensors for the noise and vibration signals, placed respectively close to and on the mechanisms, are a piezoelectric pre-amplified microphone PCB 378B02 and three piezoelectric accelerometers Dytran 3145AG (with a bandwidth of 0.5–10000 Hz and nominal sensitivity 100 mV/g).

Software

The software to run the National Instruments test bench is developed by the LabVIEW graphical environment. The main functions of the implemented algorithms are the control of the mechanism motion and the acquisition of the physical quantities: $v(t)$, $i(t)$, $n(t)$, $p(t)$, and $a(t)$.

In order to fully exploit the potentialities of the Compact RIO platform, the software is structured in the three levels that characterize the platform architecture, namely the FPGA, Real Time and PC Host. The Field Programmable Gate Array (FPGA) is an integrated circuit connected to the I/O modules that operates extremely fast (computational velocity up to 40 MHz) but is not a very flexible tool, being able to perform basic operations only. On the opposite side, the PC Host level offers a good interface with the user since it is provided with a wide library of functions, but at the price of the incapability of being able to perform deterministic, real-time and control applications due to a significantly lower computation speed. At the intermediate level is the Real Time (RT) processor that allows FPGA and PC Host to work in synergy. In this context, a proper organization of the algorithms makes it possible to optimize the software performance (in terms of accuracy and velocity).

Relative to the main tasks of this application, the algorithms implemented at the FPGA level (Fig. 4) are responsible for: (1) the generation of the Pulse Width Modulation (PWM) signals needed to move the motor according to the desired trajectory (which is computed at the PC Host level, processed by the RT processor and finally stored in the RAM memory of FPGA); (2) the implementation of the close-loop control scheme based on a popular PID controller; (3) the acquisition of the sensors' data. The experimental signals are gathered in three groups according to their nature or the required sampling frequency:

- data from the encoder (digital signals)
- data from the current-sense and voltage sensors (analog low-frequency signals needed for the calculation of the global and mechanical efficiencies)
- data from the accelerometers and the microphone (analog high-frequency signals required for the vibro-acoustic analysis).

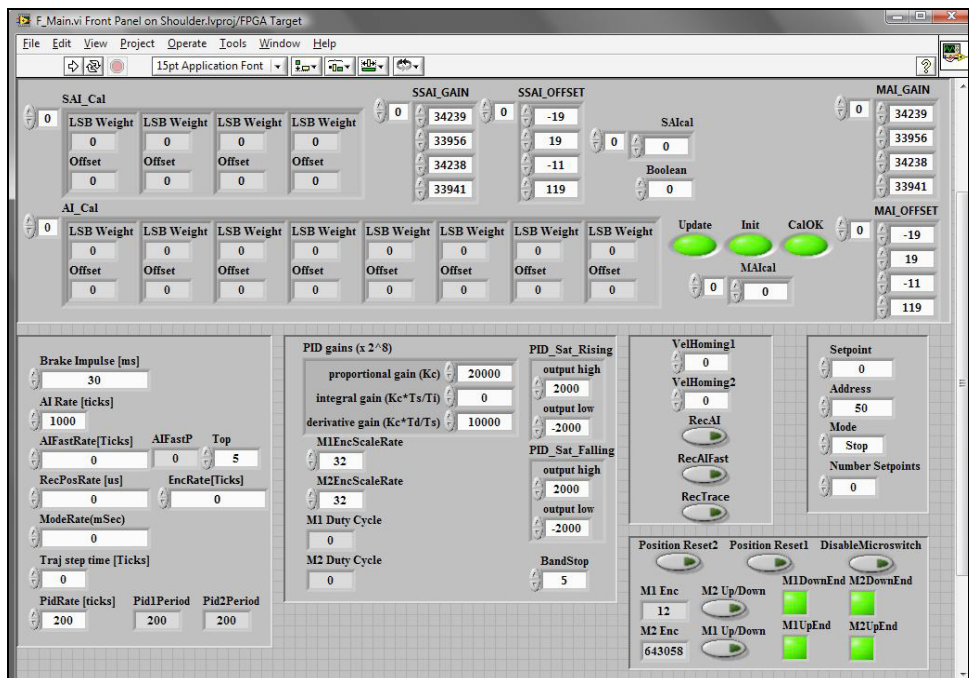


Figure 4. Front Panel of the algorithm implemented in the FPGA unit.

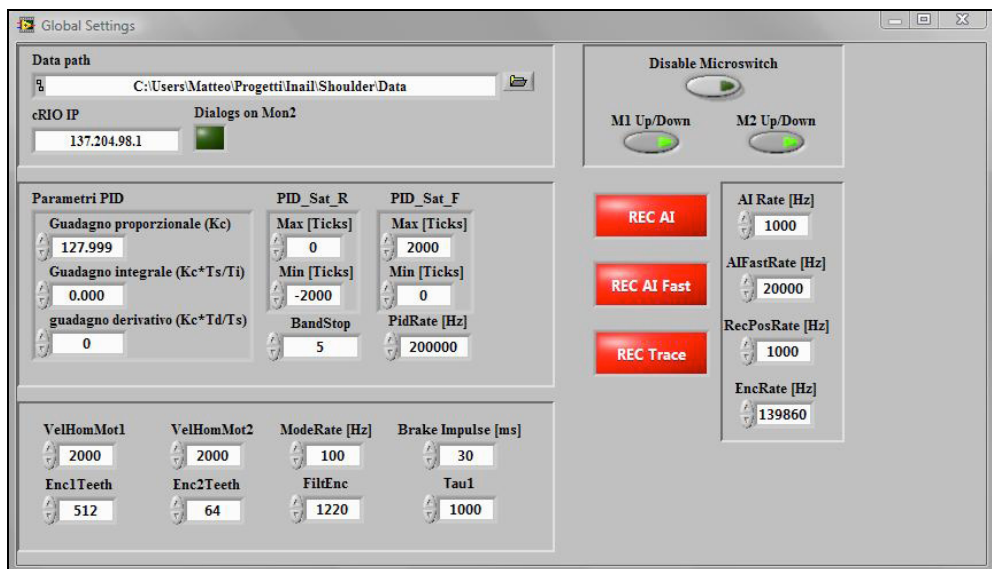


Figure 5. Front Panel of one menu available as user interface at the PC Host level.

Each group of signals is written, sample by sample, on a FIFO DMA (First-In-First-Out Direct-Memory-Access) and sent to the RT level.

The codes implemented in the RT processor are responsible for the processing of the information provided by the user at the PC Host level (e.g. reference trajectories, PID gains, data acquisition parameters...) for the correct execution of the tests. Also the conversion of the input coming from the PC Host into a format suitable for the FPGA unit and the efficient management of the acquired data coming from the FPGA and directed to the PC Host are important functions implemented at the RT level.

The user interface is implemented at the PC Host level, where a series of menus (see as an example Fig. 5) allows the configuration of the parameters needed for the motion control (e.g. reference trajectory, PID gains, encoder resolution, mechanism reduction ratio...) and the data acquisition (e.g. sets of data, sampling frequency...). At this level a number of charts are available that make it possible to check the accuracy of the actual trajectory with respect to the reference one, and to visualize the raw experimental data to verify the soundness of the performed trial. The post processing algorithms, whose definition is in progress, will be available at this level in the near future.

5. DISCUSSION

At the time of writing this paper, some results of the first application of the presented testing procedure are available. In particular the global and mechanical efficiencies of the Elbow, Shoulder-1 and Shoulder-2 mechanisms were determined and deeply discussed in [5]. In this context it is enough noting that the Elbow, which is a product rather “mature”, exhibits for all the test trials (having different payload and velocities) higher values of efficiency with respect to the two shoulder prototypes.

As far as the vibro-acoustic characterization is concerned, only preliminary results were retrieved (and actually were used for the set up of the test rig and the definition of the vibro-acoustic test procedures and algorithms). The noise level of the three mentioned prototypes was calculated (by placing the sensor 0.35 m far from the tested mechanism, consistently with the distance of the human ear from the prosthetic devices mounted on the artificial arm worn by the amputees) and are reported in Tab. 3. The data confirm that the shoulder prototypes are more noisy than the Elbow, whose higher quality is detectable also considering this aspect.

An advanced study of the vibro-acoustic data is foreseen in the next future. In particular the spectral analysis of the sound signals will be performed by computing the 1/3-octave band spectrum (that is the most popular method to investigate the sound energy distribution in the frequency domain) in order to link the noise perception to mechanical phenomena associated with the mechanisms’ functioning. To enforce this analysis and better understand its results, the analysis of mechanical vibrations will be performed as well. Apart from the statistical investigation of the acceleration signals in the time domain (Fig. 6 shows that the vibrations measured on Shoulder-2 are more severe than those of Shoulder-1, consistently with what may be concluded from Tab. 3), the spectral analysis of the data will be carried

Table 3. Sound Pressure Level of the prosthetic devices.

	Ground Noise	Elbow	Shoulder-1	Shoulder-2	Humeral Rotator
L_p [dBA]	38	48	55	64	N/A

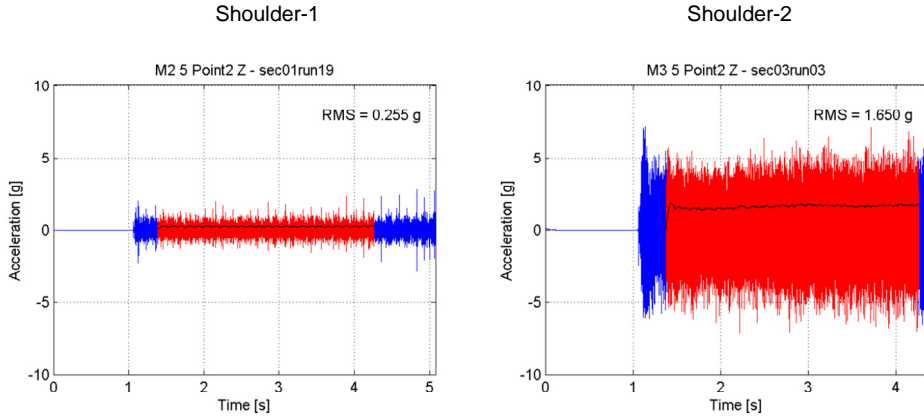


Figure 6. Acceleration signals picked-up on Shoulder-1 (left) and Shoulder-2 (right) mechanisms.

out in order to get a diagnostic tool based on the vibration frequency content [7]. Indeed, the behaviour of many mechanical components (in particular of gear systems) can be deeply investigated by associating their vibrational signatures with given mechanical phenomena, thus allowing engineers to understand how to improve the functioning of the device. For instance, in the case of gear eccentricity, the analyst can detect the fault by inspecting the spectrum of the acceleration signal (retrieved from a transducer placed close to the gear shaft bearings); in particular, this kind of fault causes modulation of the gear meshing frequency at 1x revolution of the eccentric gear (the gear meshing frequency and its harmonics are surrounded by sidebands spaced at the rotational frequency of the eccentric gear). It could be the case that the detected defect is contributing to the global sound level of the mechanism (and this could be verified by correlating the vibration spectrum to the 1/3-octave band spectrum of the noise signals). As a consequence, the design of the device (and/or the accuracy of its manufacturing/assembly) should be improved in order to guarantee the correct support of the shaft and to avoid the undesired eccentricity. The lowering of the noise level and the improvement of the mechanical and global efficiencies are expected as a final result of this procedure.

6. CONCLUSIONS

This paper presented the campaign of experimental tests designed to measure the efficiency and to investigate the vibro-acoustic behaviour of prototypes of prosthetic articulations for myoelectric artificial upper limbs. The test bench, the test procedures and the software architecture are described, and the motivations of this activity is presented. The work is still in progress: in particular, the implementation of the algorithms for the vibro-acoustic data processing (both in the time and the frequency domains) is at its final step. The application of this procedure for the complete experimental characterization of four prototypes (namely the Elbow, Shoulder-1, Shoulder-2, and the Humeral Rotator joints) will be the following phase.

Acknowledgements

The INAIL Prosthesis Center is gratefully acknowledged for financial support, the use of facilities and technical collaboration.

REFERENCES

- [1] Weir, R.F.ff., 2003. "Design of artificial arms and hands for prosthetic applications". In *Standard Handbook of Biomedical Engineering and Design*. Kutz, M. Editor, Mc Graw-Hill, New York (USA), Chapter 32, pp. 32.1 – 32.60
- [2] <http://www.inail-starter.org/Downloads.html>
- [3] Troncossi M., Gruppioni E., Chiossi M., Cutti A.G., Davalli A., Parenti-Castelli V., 2009. "A Novel Electromechanical Shoulder Articulation for Upper-Limb Prostheses: from the Design to the First Clinical Application". *Journal of Prosthetics and Orthotics*, **21**(2), pp. 79 – 90
- [4] Caminati R., 2009. "Feasibility Study, Design and Prototype Manufacturing of a New Powered Humeral Rotator for Upper Limb Myoelectric Prostheses", Ph.D. Dissertation (written in Italian), University of Bologna (Italy)
- [5] Troncossi M., Parenti-Castelli V., Chiossi M., Davalli A., 2007. "Experimental characterization of prosthetic mechanisms with one-degree of freedom", *Proceedings of XVIII AIMETA Congress*, September 11-14 2007, Brescia (Italy)
- [6] International standard IEC 61672:2003: Electroacoustics - Sound level meters.
- [7] Wowk V., 1991, *Machinery Vibration: Measurements and Analysis*. Mc Graw-Hill Inc., New York (USA).

SENSITIVITY ANALYSIS OF 1-DOF EQUIVALENT MECHANISMS FOR THE KINEMATIC MODELLING OF THE HUMAN KNEE

Nicola Sancisi

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: nicola.sancisi@unibo.it*

Diego Zannoli

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: diego.zannoli2@unibo.it*

Vincenzo Parenti Castelli

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: vincenzo.parenti@unibo.it*

Abstract. *Prosthesis and orthosis design requires mathematical models that simulate the motion of human joints. These models should be topologically simple, should reproduce the anatomical motion with high accuracy and should have a high numerical stability. As for the knee joint, one-degree-of-freedom mechanisms have been formerly proposed in order to achieve this target; in particular, two mechanisms proved to be interesting, both of them having their own peculiarities. The first one replicates the anatomical motion very well and the second one is very simple from a mechanical point of view, although slightly less accurate. Moreover, the large difference between the time required by the identification procedure of the two mechanisms and the different number of singularity problems led to suppose that the second model could also have a higher numerical stability than the first one. In this context, high numerical stability is associated with three different properties of mechanisms: low sensitivity to geometrical parameter variation, low disposition to generate singularity problems and high regularity of motion. The numerical stability of both mechanisms is quantitatively assessed in this paper to verify the preliminary supposition, by analyzing the three properties mentioned above. The results proved that, as hypothesized, the second model performs better than the first one from this point of view. This characteristic, combined with the mechanical simplicity and the acceptable accuracy, makes the second model very interesting in many practical applications, such as exoskeleton, prosthesis and orthosis design.*

Keywords: *Knee, equivalent mechanisms, numerical stability, sensitivity, singularities*

1. INTRODUCTION

The knee is one of the most studied joints of the human body because of its importance in daily life and because of the high number of injuries and diseases involving this joint. In particular, in rehabilitation field, it is very important to have mathematical models that replicate the anatomical motion of the knee. These models can be used in fact as tools to design prostheses and orthoses that replicate the natural motion of the joint with high accuracy. Two approaches have been proposed in the literature for joint modelling: a simultaneous and a sequential approach as evidenced in [1]. According to the former [2–6], the joint is modelled as a set of elastic and dumping elements whose parameters are computed all at once: in this way, kinematic and static analyses are performed at the same time, by solving motion and force equations together. According to the latter [1, 7–11], the static and the kinematic problems are solved sequentially, by considering the motion of the knee in virtually unloaded conditions (passive motion) at first, for the static and dynamic problems to be solved afterwards. The sequential approach is based on the experimental evidence that the passive motion of the knee is guided by few anatomical elements, it exhibits one degree of freedom (1DOF) [1, 9–11] and can be reproduced by a corresponding equivalent rigid spatial mechanism (ESM) [10–14]. Thus, the kinematic problem can be solved at a first step, by considering the articular structures that guide the joint passive motion and by defining the ESM that reproduces this motion. Then, at subsequent steps, the elements of this model, that have been considered as rigid until this point, are considered elastic and the remaining articular structures are added to the model, thus allowing static and dynamic problems to be solved. This approach proved to be very efficient and provides a procedure that, step by step, generates models with different grades of accuracy, thus applicable in different fields.

Several ESMs that replicate the passive motion of the knee have been proposed [10–14]. In particular, two of them proved to be interesting from different points of view. The first one [12], hereinafter called ESM2 for consistency with the previous literature [11], is based on this anatomical evidence: one fibre within the anterior cruciate (ACL), one within the posterior cruciate (PCL) and one within the medial collateral (MCL) ligaments are almost isometric during passive motion [15–17] and thus are modelled as binary links of constant length which are connected at their end to the femur and to the tibia through spherical pairs. Moreover, both the medial and lateral condyles are almost spherical and thus are modelled as two spheres (one on the femur, the other on the tibia) that are constrained to remain in contact throughout the flexion arc. The second model, which is a spherical parallel mechanism (SPM) [14], is based both on anatomical and kinematic evidence. Indeed, since the instantaneous screw axes of the femur and tibia relative motion pass all near a single point (hereinafter called pivot point) [18, 19], this motion can be considered approximately spherical. According to this hypothesis, the femur and tibia were connected by a spherical pair centred at the pivot point. Moreover one fibre within the ACL and one fibre within the PCL are considered isometric and are modelled as binary links of constant length as in the previous case.

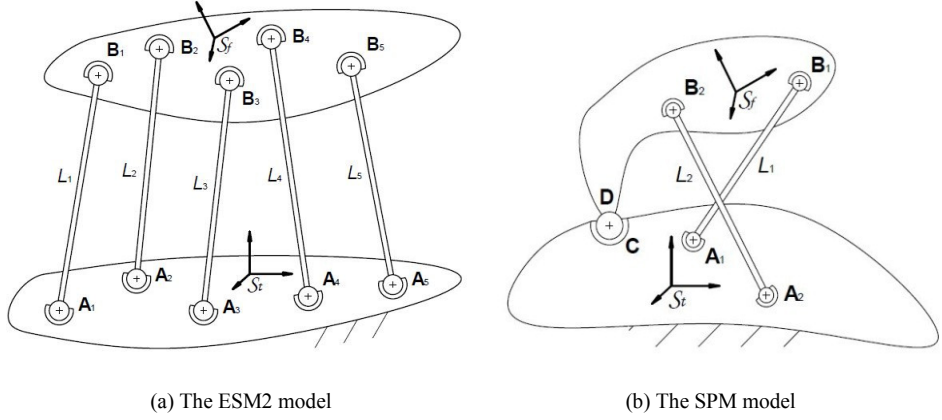


Figure 1. The two kinematic models of the knee joint. The geometrical parameters are also represented.

The ESM2 model proved to be very accurate in reproducing the experimental passive motion; the SPM model was slightly less accurate but it is equally interesting for practical applications because of its mechanical simplicity [14]. In this study, numerical stability was estimated for both models and for four relevant specimens, and the results are here presented and discussed. In this context, a high numerical stability is associated with three different properties of mechanisms: low sensitivity to geometrical parameter variation, low disposition to generate singularity problems and high regularity of motion. These characteristics, if correctly assessed and verified, would contribute to make the SPM model very interesting in practical applications like musculo-skeletal modelling and exoskeleton, prosthesis and orthosis design.

2. METHODS

Kinematic Models

The equivalent 1DOF mechanism ESM2 was defined as reported in [11, 12]. The femur and tibia were represented as rigid bodies connected by five rigid binary links through ten spherical pairs, as depicted in Fig. (1a). Three links respectively identify the ACL, PCL and MCL. The other two links replace the two pairs of spheres which approximate the condylar surfaces and are constrained to remain in contact during motion. One anatomical reference frame, S_f , was embedded in the femur, another one, S_t , in the tibia, to describe the relative motion between these two bones. The closure equations of the mechanism may be represented by the following five equations:

$$\|A_i - RB_i - P\| = L_i \quad (i = 1, \dots, 5), \quad (1)$$

where A_i ($i = 1, \dots, 5$) is the connection point, represented in S_t , between the tibia and the i -th rigid link; B_i ($i = 1, \dots, 5$) is the connection point, represented in S_f , between the

femur and the i -th rigid link; L_i ($i = 1, \dots, 5$) is the length of the i -th rigid link; R is the 3×3 rotation matrix that transforms vector components from S_f to S_t ; \mathbf{P} is the position vector of the origin of S_f in S_t . This model is defined by 35 independent geometrical parameters. For each value assumed by the independent variable, i.e. the flexion angle, Eqn. (1) produces corresponding values of the other pose variables, namely three displacements and two rotations, hereinafter called secondary variables.

The equivalent 1DOF mechanism SPM was defined as reported in [14]. The femur and tibia were represented as rigid bodies connected to each other by a spherical pair; moreover, the two bones are also interconnected through two binary links via spherical pairs (Fig. (1b)). The two rigid binary links model the ACL and PCL. The centre of the spherical pair connecting the femur and tibia is represented by the point \mathbf{C} in S_t and by the point \mathbf{D} in S_f . The closure equations of this mechanism can be written as follows:

$$\begin{cases} \|\mathbf{A}_i - R\mathbf{B}_i - \mathbf{P}\| = L_i & (i = 1, 2) \\ \mathbf{C} - R\mathbf{D} - \mathbf{P} = \mathbf{0} \end{cases}, \quad (2)$$

where \mathbf{A}_i ($i = 1, 2$) is the centre of the spherical pair, represented in S_t , connecting the tibia and the i -th binary rigid link; \mathbf{B}_i ($i = 1, 2$) is the centre of the spherical pair represented in S_f , connecting the femur and the i -th binary rigid link; L_i ($i = 1, 2$) is the length of the i -th rigid link; R and \mathbf{P} have the same meaning as in Eqn. (1). This model is defined by 20 geometrical parameters, even though only 16 are independent. As in the previous case, for each value assumed by the flexion angle, Eqn. (2) produces corresponding values of secondary variables.

Data Acquisition and Pre-processing

Reference geometric and kinematic data for the definition of ESM2 and SPM models were obtained from four fresh-frozen lower limbs, by following a standard experimental procedure [11]. Substantially, for each specimen, two reference frames (defined by trackers fixed to the bones) were defined, the first one in the femur, the second one in the tibia, and the experimental passive motion of the joint was recorded and digitized. The motion was obtained as a set of poses, each representing the position and orientation of the two trackers with respect to a laboratory reference frame. The joint was then disarticulated and the anatomical surfaces were also digitized, together with the ligament insertion areas.

Four SPM models and four ESM2 models, each relevant to a particular specimen, were obtained from these experimental data and then were identified by an optimization procedure. The two anatomical reference frames S_f and S_t were defined according to [11, 14]. These reference frames had the x axis oriented anteriorly, the y axis oriented proximally and the z axis oriented from left to right. The origin of S_t was the centre of the tibial plateau, i.e. the deepest point in the sulcus between the medial and lateral tibial intercondylar tubercles; the origin of S_f was the mid-point between the two epicondyles. Relative knee joint rotations, i.e. flexion(+)/extension(-), ab(+)/adduction(-), in(+)/external(-) rotations, were determined according to [20], resulting in angles α , β and γ

respectively. Offsets were added so that all these angles were equal to zero at full extension. Relative translations of the femur with respect to the tibia were measured by the displacements of a point embedded in S_f , that at full extension was coincident with the origin of S_t . In particular, displacement along the x, y, z axes of S_t resulted respectively in anterior(+)/posterior(-), proximal(+)/distal(-), lateral(+)/medial(-) displacements and were equal to zero at full extension.

A first estimate of geometrical parameters was obtained by processing the anatomical surfaces which were digitized during the experimental procedure. In particular, the points A_i and B_i ($i=1,2,3$) of each ESM2 model and A_i and B_i ($i=1,2$) of each SPM model were chosen as the pair of points in the relevant ligament insertion areas whose mutual distance had the smallest variation during motion. Furthermore, the first estimate of the points **C** and **D** in SPM coincided with the pivot point of the motion, calculated by means of a standard least-square procedure [21] as the point featuring the minimum mean distance from the instantaneous screw axes of the femur and tibia relative motion. Finally, the points A_i and B_i ($i=4,5$) that defines the condylar surfaces in ESM2 were obtained by approximating the condyles by best-fitting spheres.

The geometrical parameter first estimate was refined by means of an optimization procedure in order to calculate the parameters which best approximate the experimental data. An objective function that estimates the error between the calculated and the experimental motion was defined as follows:

$$\begin{cases} f = \sum_{j=1}^5 \sum_{i=1}^n \frac{(q_{ij} - q_{ij}^*)^2}{q_{jd}^2} & \text{if closure succeeded} \\ f = X & \text{otherwise} \end{cases} \quad (3)$$

where q_{ij} and q_{ij}^* are the calculated and the experimental values of the j-th secondary variable ($j=1,...,5$) at the i-th flexion angle ($i=1,...,n$) respectively, n is the number of flexion angles, q_{jd} is the excursion of the j-th secondary variable; if the closure equations could not be solved at all flexion angles, an arbitrary high value X was assigned to the objective function. This function was hence minimized by the Newton-Raphson algorithm: the minimum identifies the particular set of geometrical parameters which define the optimum mechanism. The optimization domain was properly bounded in order to synthesize a final geometry not too different from the first estimate. At the end of this procedure, four optimized SPM models and four optimized ESM2 models were obtained.

Numerical Stability Analysis

Numerical stability of the defined mechanisms was analyzed in several respects. Firstly, sensitivity to infinitesimal variation of geometrical parameters was evaluated. The closure equations can be written in the form

$$\mathbf{g}(\mathbf{q}, \alpha, \mathbf{p}) = \mathbf{0}, \quad (4)$$

where \mathbf{g} are the closure equations, α is the flexion angle (the independent variable); \mathbf{q} and \mathbf{p} are vectors containing the secondary variables and the geometrical parameters respectively. The length m of \mathbf{p} is different for the two models, but the same symbol is used for the ESM2 and SPM models, in order to simplify the notation. If the flexion angle α is fixed, differentiation of Eqn. (4) yields

$$\mathbf{M}\delta\mathbf{q} + \mathbf{N}\delta\mathbf{p} = \mathbf{0}, \quad (5)$$

where $\delta\mathbf{p}$ represents a general infinitesimal variation of the geometrical parameters, $\delta\mathbf{q}$ represents the respective variation of the secondary variables, \mathbf{M} and \mathbf{N} are matrices containing partial derivatives of \mathbf{g} with respect to the secondary variables and to the geometrical parameters respectively. Both \mathbf{M} and \mathbf{N} depend on α . If \mathbf{M} is not singular, Eqn. (5) yields

$$\delta\mathbf{q} = -\mathbf{M}^{-1}\mathbf{N}\delta\mathbf{p} = \mathbf{H}\delta\mathbf{p}, \quad (6)$$

where \mathbf{H} is a $5 \times m$ matrix that relates the infinitesimal variation of geometrical parameters to the corresponding variation of the secondary variables [22]. In particular, each element h_{jk} of \mathbf{H} is the partial derivative of the j -th secondary variable with respect to the k -th geometrical parameter ($k=1,...,m$). Matrix \mathbf{H} was computed at all flexion angles for each optimized model.

The analysis of elements h_{jk} makes it possible to determine the sensitivity of each secondary variable to a small variation of each geometrical parameter at the given flexion angles. However, it is extremely useful and convenient to define an index which describes the sensitivity of a mechanism from a global point of view. Several properties of matrix singular values can be used to achieve this target. Firstly, matrix \mathbf{H} was made dimensionally homogeneous by dividing each element h_{jk} by the excursion of the corresponding j -th secondary variable. In other words, each row j of \mathbf{H} was divided by the corresponding value q_{jd} . In this way, a new matrix \mathbf{H}_N was obtained which satisfies the relation

$$\delta\mathbf{q}_N = \mathbf{H}_N\delta\mathbf{p}, \quad (7)$$

where $\delta\mathbf{q}_N$ is a vector that collects the relative variations of secondary variables with respect to their corresponding excursions. The elements of $\delta\mathbf{q}_N$ will be called secondary variables in the following for the sake of clarity and brevity. Matrix \mathbf{H}_N , like all rectangular matrices, can take the form

$$\mathbf{H}_N = \mathbf{U}\Sigma\mathbf{V}, \quad (8)$$

where \mathbf{U} and \mathbf{V} are respectively a 5×5 and a $m \times m$ unitary matrix and Σ is a $5 \times m$ diagonal matrix, with nonnegative elements on the diagonal. Such a factorization is the singular

value decomposition of H_N and the elements on the diagonal of matrix Σ are known as the singular values of H_N . The linear transformation identified by matrix H_N can be thus decomposed in a series of three sub-transformations: a rotation V in the m -dimensional space of geometrical parameter variations, a pure dilatation Σ that transforms from this m -dimensional space to the 5-dimensional space of secondary variable variations and a pure rotation U in this 5-dimensional space. Each singular value of matrix H_N is thus the ratio between the modulus of the vector $\delta \mathbf{q}_N$ and the modulus of the vector $\delta \mathbf{p}$ along a particular main direction of the transformation identified by matrix H_N , i.e.

$$s_j = \frac{\|\delta \mathbf{q}_N\|}{\|\delta \mathbf{p}\|}, \quad (9)$$

where s_j ($j=1,\dots,5$) is the j -th singular value of matrix H_N . For $\|\delta \mathbf{p}\| = 1$, Eqn. (9) can be written as

$$s_j = \sqrt{5} \sqrt{\frac{\delta q_{N1}^2 + \delta q_{N2}^2 + \delta q_{N3}^2 + \delta q_{N4}^2 + \delta q_{N5}^2}{5}}. \quad (10)$$

Accordingly, along each principal direction, the relevant singular value is proportional to the root mean square variation of the secondary variables.

In particular, the maximum singular value $s_{j_{\max}}$ is proportional to the maximum root mean square variation of the secondary variables caused by a unit error vector $\delta \mathbf{p}$ of geometrical parameters. Thus, $s_{j_{\max}}$ provides an important information about the sensitivity of the mechanism and was used as a first global sensitivity index I_1 . However, the other singular values also have an important role in describing sensitivity because they refer to the other principal directions of H_N . Another global sensitivity index, hereinafter called I_2 , was hence defined as the product of the five singular values. This index is proportional to the volume of the ellipsoid in the 5-dimensional space of secondary variable variations, that is the transformed of a unitary radius sphere in the m -dimensional space of geometrical parameter variations. These indices provide complementary information about the sensitivity of a mechanism: I_1 describes the maximum sensitivity of the mechanism, while I_2 describes the average sensitivity by considering all principal directions. Indices I_1 and I_2 were evaluated for each optimized model and at all flexion angles.

Sensitivity to relatively large variations of geometrical parameters was also analyzed. For each optimized model, geometrical parameters were modified by a random error chosen in a way that modified insertion points and lengths were distant no more than 1 mm from the optimal ones. For this modified mechanisms, the closure equations were solved at all flexion angles and the corresponding secondary variables were computed. If the closure equations could not be solved at all flexion angles, another random modification of geometrical parameters was produced. This procedure was repeated 3000 times for each

optimized model. Thus, 3000 mechanisms affected by random geometrical errors were obtained on the whole for each optimized mechanism, and the global mean square error of the secondary variables was calculated as follows:

$$e = \frac{1}{3000 \cdot 5 \cdot n} \sum_{l=1}^{3000} \sum_{j=1}^5 \sum_{i=1}^n \frac{(q_{lij}^e - q_{ji}^{opt})^2}{q_{jd}^2}, \quad (11)$$

where q_{lij}^e and q_{ji}^{opt} are the values of the j -th secondary variable at the i -th flexion angle obtained respectively from the l -th mechanism affected by error and from the corresponding optimized mechanism. The total number v of mechanisms for which the closure did not succeed at all flexion angles was also computed and was used as an index to estimate the tendency of a model to generate singularity problems.

Finally, the first derivatives of the secondary variables with respect to the flexion angle α were computed for the 3000 mechanisms affected by random errors and were compared with the corresponding derivatives evaluated for the optimized mechanisms. This comparison was carried out by a mean square error e_{der} defined like in Eqn. (11) by replacing the secondary variables with their derivatives. This index provides further information regarding the sensitivity of equivalent mechanisms since it quantifies how the derivatives of secondary variables are sensitive to large geometrical parameter variations. This aspect is directly related to the regularity of motion when geometrical errors affect the mechanism.

3. RESULTS AND DISCUSSION

The optimization procedure was carried out on a standard platform with a 2.8 GHz processor and 6 GB of RAM; it took about one hour for ESM2 model and about five minutes for SPM model for each specimen. The pre-optimization process that was required to overcome singularity problems took several hours for ESM2 model and about five minutes for SPM model. This high difference between the two mechanisms is due to the greater mechanical simplicity of the SPM model with respect to the ESM2 model.

Accuracy analysis and relevant results of both ESM2 and SPM were presented in previous investigations [11, 12, 14] and are not reported here. In summary, model ESM2 replicates the experimental motion very well. Model SPM is slightly less accurate but the results still remain within the range of acceptance for many application fields such as exoskeleton, prosthesis and orthosis design, and musculo-skeletal modelling.

In this paper, numerical stability was evaluated and the results are here reported. The elements of the $5 \times m$ matrix H_N are shown in Fig. (2) for one representative specimen. Similar results were obtained for the other specimens. The left column refers to ESM2 model and the right column refers to SPM model while the j -th row ($j=1, \dots, 5$) reports h_{jk} ($k=1, \dots, m$) elements of H_N versus the flexion angle α ; each colour refers to a different geometrical parameter k . This figure shows that the sensitivity to infinitesimal geometrical

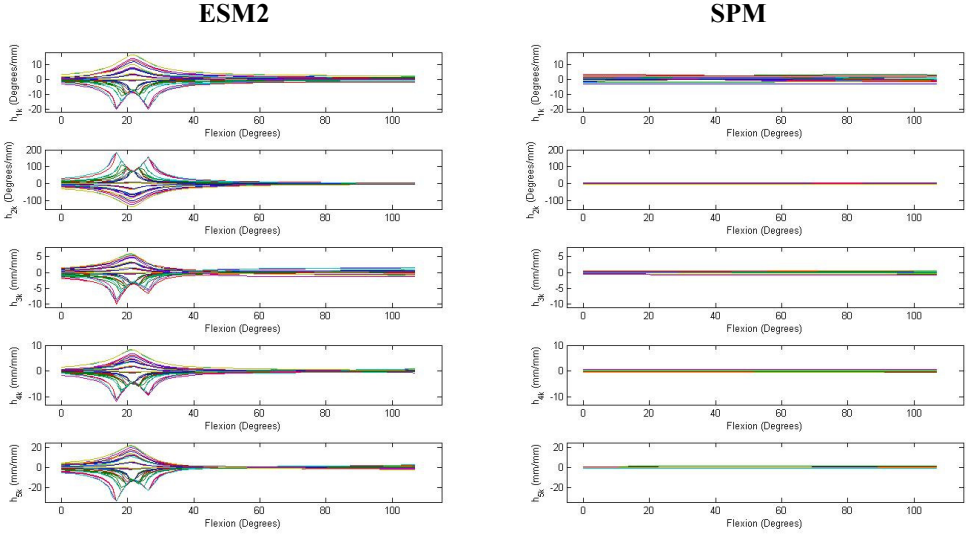


Figure 2. Sensitivity to small variations of the geometrical parameters of ESM2 and SPM models.

parameter variations is much higher in the ESM2 model than in the SPM model. Furthermore, both global sensitivity indices I_1 and I_2 are represented in Fig. (3) for the same specimen of Fig. (2). In particular, the first row shows the index I_1 in thick blue line. The other four singular values of matrix H_N are reported in thin lines for completeness. The second row shows the index I_2 versus the flexion angle α . These indices confirm the higher sensitivity of the ESM2 model with respect to the SPM model. In particular, for the specimen considered in the figures, the index I_1 is much higher for ESM2 model in the first half of the flexion arc and has about the same value for both models in the second half. The index I_2 instead is much higher for ESM2 model throughout the flexion arc, and the difference still increases in the first half of the motion. The fact that these indices confirm the trend that has been shown in Fig. (2) is a partial verification of their capacity to represent and summarize the mechanism sensitivity to geometrical parameter variation.

The results that correspond to large variations of geometrical parameters are shown in Fig. (4) for the same specimen of the previous figures. As before, the left column refers to the ESM2 model, the right column refers to the SPM model and the j -th row shows the j -th secondary variable patterns versus the flexion angle α . The thick red lines represent the secondary variable patterns of the optimized mechanism, the thin blue lines represent the same patterns of the 3000 random modified mechanisms, the thick green line is the mean value of the 3000 blue lines and the thick green dashed lines define the relevant standard deviation. The higher dispersion of the blue lines and the corresponding higher standard deviation computed from ESM2 model confirm its higher sensitivity with respect to SPM model even in the presence of large geometrical parameter variations. Moreover, for the SPM model, the mean value of each secondary variable (green lines) coincides approximately with the corresponding value of the optimized mechanism (red lines); on the

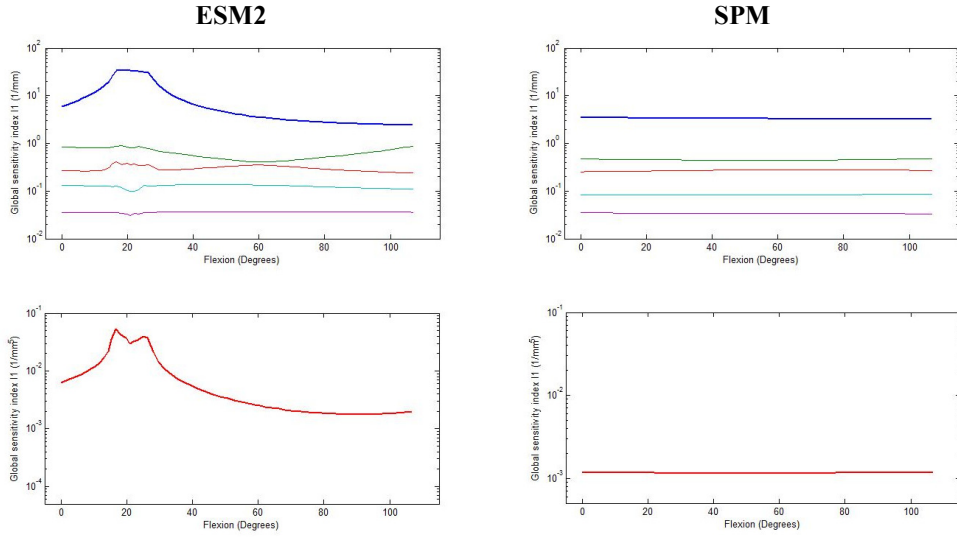


Figure 3. Global sensitivity indices I_1 and I_2 evaluated for ESM2 and SPM models.

contrary, this does not occur for the ESM2 model, for which these values are very different. This fact shows the strong nonlinearity of the ESM2 model with respect to the SPM model. Furthermore, the bifurcations in blue patterns of the ESM2 model indicate its tendency to converge to different closure configurations; this aspect reveals the higher tendency of the ESM2 model to generate singularity problems.

The global weighted mean square error e , as defined in Eqn. (11), the corresponding error e_{der} related to the derivatives of secondary variables, and the total number v of mechanisms for which the closure did not succeed at all flexion angles are reported in Tab. (1) for both models and for all specimens. These indices are always higher for the ESM2 model than for the SPM model. This evidence confirms the higher numerical stability of SPM with respect to ESM2: SPM generates a more stable and regular motion and has a lower tendency to generate singularity problems.

Table 1. Global error indices applied to the secondary coordinate patterns (e) and to their first derivative (e_{der}), and the total number v of mechanisms for which closure did not succeed at all flexion angles.

		Specimen 1	Specimen 2	Specimen 3	Specimen 4
e	ESM2	0.1075	16.9209	2.3964	3.1947
	SPM	0.0861	0.0620	0.2955	0.5259
e_{der}	ESM2	0.0488	4.1510	0.0350	0.4054
	SPM	0.0045	0.0049	0.0152	0.0451
v	ESM2	172	2430	12235	8557
	SPM	0	0	854	0

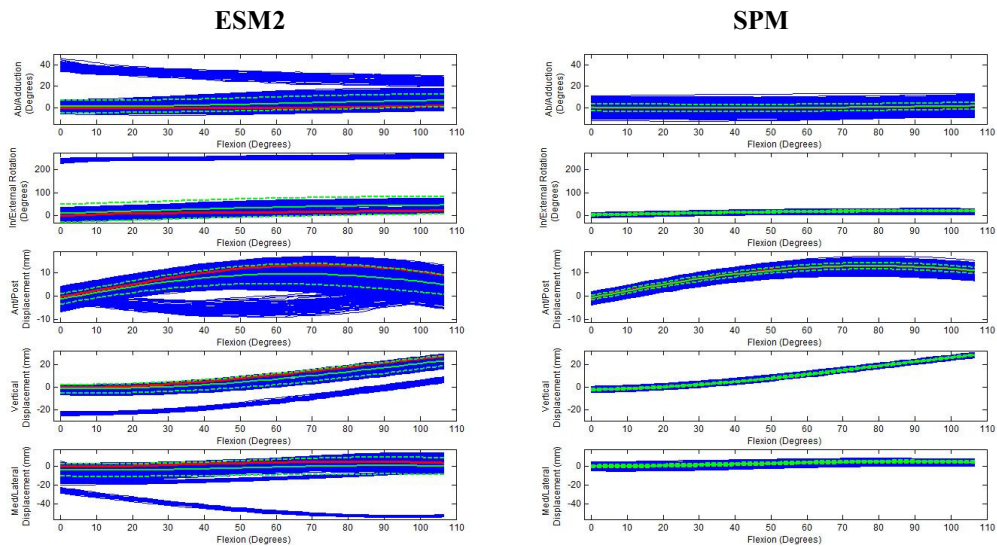


Figure 4. Sensitivity to large variations of the geometrical parameters of ESM2 and SPM models.

4. CONCLUSIONS

Mathematical models that simulate the passive motion of human joints are very useful in rehabilitation field and prosthesis and orthosis design. These models should be topologically simple, should be accurate in describing anatomical motion and should have a high numerical stability.

In this paper, the numerical stability of two classes of mechanisms for the kinematic modelling of the knee joint (ESM2 and SPM models) was analyzed. The model accuracy was studied in previous papers: ESM2 reproduces the experimental motion very well; the SPM is slightly less accurate but, despite its greater mechanical simplicity, it has a good accuracy for many practical applications. According to the results of this paper, the SPM model proves also to have a higher numerical stability than the ESM2, showing a lower sensitivity to geometrical parameter variations, a lower disposition to generate singularity problems and a higher regularity of motion. Thus, good accuracy, mechanical simplicity and high numerical stability make the SPM model very interesting in practical applications like exoskeleton, prosthesis and orthosis design, and musculo-skeletal modelling.

Acknowledgements

The financial support of AERTECH-LAB project, funded by Regione Emilia-Romagna, and MIUR is gratefully acknowledged. The authors also thank Dr. Alberto Leardini for the facilities and Dr. Claudio Belvedere for the support during experimental sessions.

REFERENCES

- [1] V. Parenti-Castelli, A. Leardini, R. Di Gregorio, and J. J. O'Connor, 2004. "On the modeling of passive motion of the human knee joint by means of equivalent planar and spatial parallel mechanisms". *Autonomous Robots*, **16**(2), pp. 219–232.
- [2] M. S. Hefzy and T. D. V. Cooke, 1996. "Review of knee models: 1996 update". *Applied Mechanics Reviews*, **49**(10-2), pp. 187–193.
- [3] M. G. Pandy, K. Sasaki, and S. Kim, 1997. "A three-dimensional musculoskeletal model of the human knee joint. Part 1: theoretical construction". *Computer Methods in Biomechanics and Biomedical Engineering*, **1**, pp. 87–108.
- [4] M. G. Pandy, 2001. "Computer modeling and simulation of human movement". *Annual Review of Biomedical Engineering*, **3**, pp. 245–273.
- [5] K. F. T. Olanlokun and D. P. M. Wills, 2002. "A spatial model of the knee for the preoperative planning of knee surgery". *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine*, **216**(H1), pp. 63–75.
- [6] K. E. Moglo and A. Shirazi-Adl, 2005. "Cruciate coupling and screw-home mechanism in passive knee joint during extension-flexion". *Journal of Biomechanics*, **38**, pp. 1075–1083.
- [7] N. Sancisi and V. Parenti-Castelli, 2008. "A sequential approach for modelling knee joint stiffness". In *proceedings of Romansy 2008*, Tokyo, Japan, July 05-09, pp. 1–8.
- [8] R. Franci, V. Parenti-Castelli, and N. Sancisi, 2009. "A three-step procedure for the modelling of human diarthrodial joints". *International Journal of Mechanics and Control*, **10**(1), pp. 3–12.
- [9] J. J. O'Connor, T. W. Lu, D. R. Wilson, J. D. Feikes, and A. Leardini, 1998. "Diarthrodial joints-kinematic pairs, mechanisms or flexible structures?". *Computer Methods in Biomechanics and Biomedical Engineering*, **1**(2), pp. 123–150.
- [10] D. R. Wilson and J. J. O'Connor, 1997. "A three-dimensional geometric model of the knee for the study of joint forces in gait". *Gait and Posture*, **5**, pp. 108–115.
- [11] A. Ottoboni, V. Parenti-Castelli, N. Sancisi, C. Belvedere, and A. Leardini, 2010. "Articular surface approximation in equivalent spatial parallel mechanism models of the human knee joint". *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. DOI: 10.1243/09544119JEIM684, pp. 1–11.
- [12] V. Parenti-Castelli and R. Di Gregorio, 2000. "Parallel mechanisms applied to the human knee passive motion simulation". In *Advances in Robot Kinematics*, J. Lenarcic and M. Stanisic, Eds. Kluwer Academic.
- [13] R. Di Gregorio and V. Parenti-Castelli, 2003. "A spatial mechanism with higher pairs for modelling the human knee joint". In *Journal of Biomechanical Engineering*, **125**(2), pp. 232–237.
- [14] N. Sancisi and V. Parenti-Castelli, 2010. "A 1-dof parallel spherical wrist for the modelling of the knee passive motion". *Mechanism and Machine Theory*, **45**, pp. 658–665.
- [15] D. R. Wilson, J. D. Feikes, and J. J. O'Connor, 1998. "Ligament and Articular Contact Guide Passive Knee Flexion". *Journal of Biomechanics*, **31**, pp. 1127–1136.
- [16] S. K. Fuss, 1989. "Anatomy of the Cruciate Ligaments and Their Function in Extension and Flexion of the Human Knee Joint". *American Journal of Anatomy*, **184**, pp. 165–176.
- [17] J. S. Rovick, J. D. Reuben, R. J. Schrager, and P. S. Walker, 1991. "Relation Between Knee Motion and Ligament Length Patterns". *Clinical Biomechanics*, **6**(4), pp. 23–220.

- [18] L. Blankevoort, R. Huiskes, and A. D. Lange, 1990. "Helical axes of passive joint motions". *Journal of Biomechanics*, **23**, pp. 1219–1229.
- [19] F. K. Fuss, 1993. "Helical axis surface of the knee joint". In Proceedings of 14th Congress of the International Society of Biomechanics, Paris, France, July 4–8.
- [20] E. S. Grood and W. J. Suntay, 1983. "A joint coordinate system for the clinical description of three-dimensional motions: application to the knee". *Journal of Biomechanical Engineering*, **105**, pp. 136–144.
- [21] H. J. Woltring, 1990. "Data processing and error analysis". In *Biomechanics of Human Movement Applications in Rehabilitation Sports and Ergonomics*, N. Berme and A. Cappozzo, Eds. Bertec Corporation, pp. 203–307.
- [22] R. Di Gregorio and V. Parenti-Castelli, 2005. "Systematic sensitivity analysis of spatial one-dof models of diarthrodial joints". In Proceedings of IDETC/CIE 2005, Long Beach, California, USA, September 24–28, pp. 1–9.

APPLICAZIONI STORICHE DELLE CAMME

Umberto Meneghetti

DIEM – Dipartimento di Ingegneria delle Costruzioni meccaniche, nucleari, aeronautiche e di Metallurgia,

Università di Bologna, Italia

E-mail: umberto.meneghetti@mail.ing.unibo.it

Sommario. *Le camme sono organi meccanici ben noti e sono oggi largamente impiegate, per esempio nei motori a combustione interna, nelle macchine confezionatrici e in molti altri campi. Anche nel passato esse ebbero molte interessanti applicazioni, svolgendo un ruolo fondamentale nello sviluppo di settori quali la siderurgia, la metallurgia, l'industria della carta, l'industria tessile, e così via. Semplici alberi a camme, sotto forma di rulli o cilindri a perni, si trovano impiegati per azionare strumenti musicali quali gli organi meccanici e le scatole musicali, mentre il funzionamento degli automi del Settecento si basa sull'impiego di camme molto evolute e sofisticate. Esempi più recenti, ma ormai storici, di applicazioni delle camme si trovano nelle macchine da scrivere, nei calcolatori analogici, nei servomeccanismi, nei torni automatici.*

Questo lavoro illustra brevemente alcune delle applicazioni sopra citate.

Parole chiave: *alberi a camme, applicazioni storiche, automi, camme*

1. INTRODUZIONE

Secondo il Dizionario della Lingua italiana Devoto-Oli, il termine *camma* deriva dal francese *came*, e secondo il Dizionario Francese Larousse, questo dal tedesco *Kamm*, cioè *cresta*.

L'aspetto dei primi alberi a camme giustifica pienamente questa etimologia. Si tratta, infatti, di alberi muniti di grossolane sporgenze destinate ad imprimere un moto alterno ad un membro traslante od oscillante, per azionare macchine quali mantici, pompe, seghe, magli, pile e simili. La forma delle camme non evidenzia un particolare interesse per la legge di moto impressa al cedente, che deve semplicemente essere spostato da una posizione di partenza ad una di arrivo; il ritorno viene poi ottenuto per azione di una molla o, più comunemente, per effetto del peso dello stesso cedente. Queste applicazioni sono proseguite, senza modifiche essenziali, almeno fino a tutto il XIX secolo.

Anche le interessanti applicazioni ai carillon, agli organi meccanici e alle scatole musicali fanno ricorso a camme sostanzialmente simili a quelle sopra menzionate, con solo qualche piccola ulteriore caratterizzazione suggerita dal tipo di applicazione. Nella seconda metà del XVIII secolo, invece, fanno la loro comparsa camme analoghe a quelle impiegate nelle attuali applicazioni industriali, cioè camme destinate ad imprimere al cedente sofisticate leggi di moto. Da queste camme hanno preso le mosse quelle impiegate nelle macchine automatiche, nei servomeccanismi e nelle altre complesse applicazioni che si incontrano dal XX secolo fino ai nostri giorni.

Questo lavoro si propone di illustrare brevemente le principali applicazioni storiche delle camme, senza entrare nel merito delle specifiche storie dei molti, importanti settori in cui tali applicazioni si sono sviluppate.

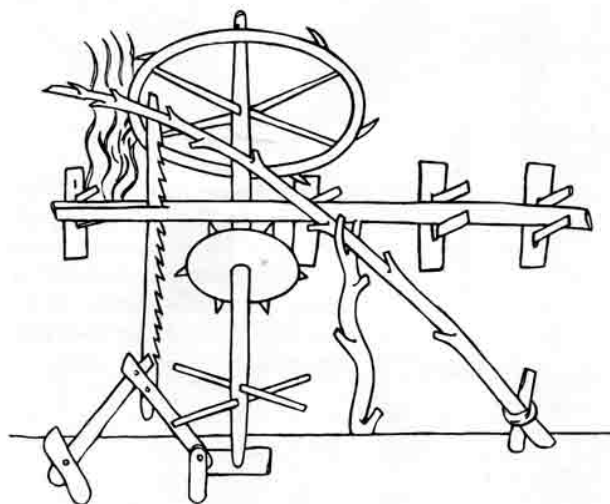


Figura 1. Sega idraulica azionata da un albero a camme, dal “Livre de portaiture” di Villard de Honnecourt, XIII secolo [1].

2. I PRIMITIVI ALBERI A CAMME

Tralasciando possibili applicazioni più antiche, come quelle di epoca ellenistica – per le quali si hanno descrizioni letterarie, ma non rappresentazioni grafiche contemporanee –, si ritiene comunemente che la prima raffigurazione di un albero a camme sia quella relativa alla sega idraulica presentata nel suo taccuino di appunti da Villard de Honnecourt [1], ingegnere-architetto francese del XIII secolo. Villard ha rappresentato lo schizzo di una sega alternativa, mossa da una ruota idraulica che trascina un albero a camme: queste impongono uno spostamento ad un membro di un sistema articolato che, a sua volta, muove la sega in un verso; il ritorno, cioè il movimento della sega nel verso opposto, è prodotto da un ramo fissato al terreno e funzionante da molla, v. Fig. 1.

Da notare che Villard non ha inventato la sega che rappresenta nel suo taccuino, ma si è limitato ad annotare una cosa che ha visto nei suoi viaggi e che gli è parsa interessante. Possiamo quindi presumere che applicazioni del genere – forse di derivazione araba –, per quanto non comunissime, fossero diffuse da tempo in Europa.

In effetti, si hanno notizie di mulini a camme per la follatura della lana presenti in Inghilterra e in Spagna già nel XII secolo.

Un’applicazione storicamente accertata è comunque quella della pila di Fabriano [2], v. Fig. 2, risalente al XIII secolo: si tratta di un albero a camme impiegato per sollevare i pestelli della pila, o mulino, in cui gli stracci venivano ridotti in poltiglia, successivamente impiegata per la fabbricazione della carta. Proprio all’impiego di questo dispositivo meccanico viene attribuita gran parte del merito della superiore qualità della carta prodotta a Fabriano.

Successivamente, e per tutto il Rinascimento e oltre, si incontrano numerosissime rappresentazioni di alberi a camme [3], [4], [5], [6], [7]. Molte di queste rappresentazioni si riferiscono ad applicazioni reali, o almeno possibili, v. Figg. 3, 5, mentre altre appaiono solo



Figura 2. Pila o Mulino di Fabriano. Vecchio albero a camme (sopra) e ricostruzione moderna della pila a magli multipli (sotto).

esercizi di fantasia, spesso, peraltro, non privi di interesse, come nel caso della camma a comando positivo rappresentata da Agostino Ramelli, v. Fig. 4.

Gli alberi a camme, nella forma primitiva della pila di Fabriano o in forme poco più evolute, sono stati sicuramente impiegati almeno fino al XIX secolo, per es., in industrie tessili, metallurgiche e siderurgiche, in pile da riso, e altro [8], [9], come attestano per es. varie rappresentazioni in manuali ed enciclopedie, v. Fig. 6.

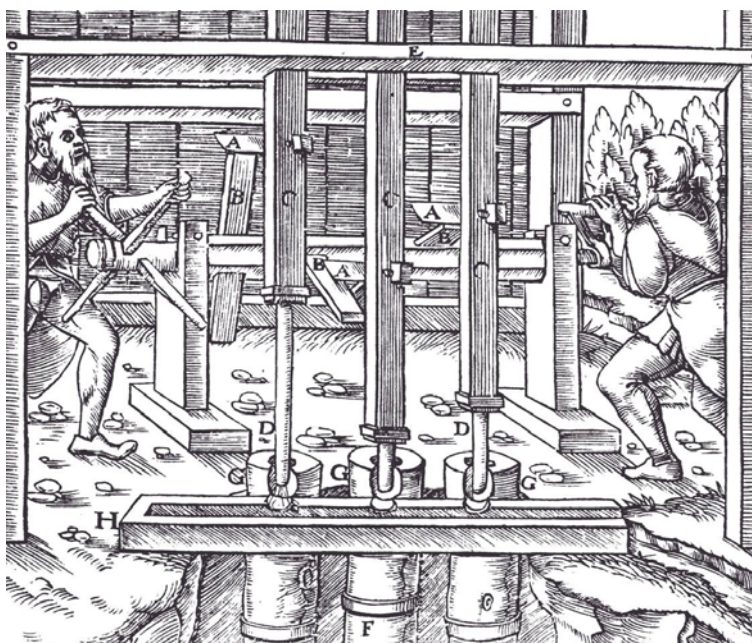
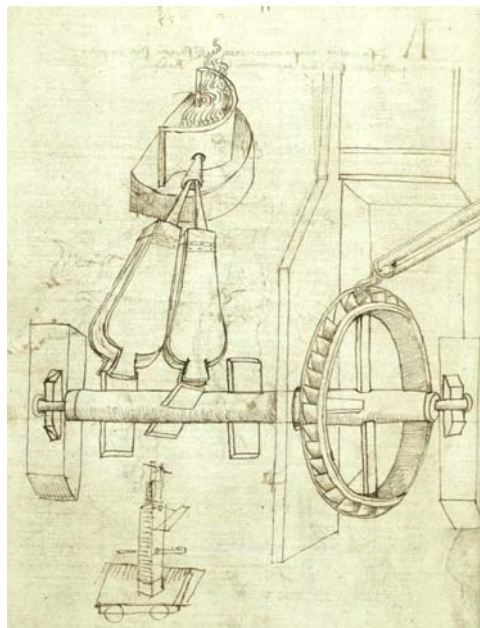


Figura 3. Esempi di alberi a camme. Sopra: Mantici, disegno del Taccola (Mariano di Jacopo), 1382-1453 circa [3]; sotto: Pompa da miniera di Giorgio Agricola, 1563 [5].

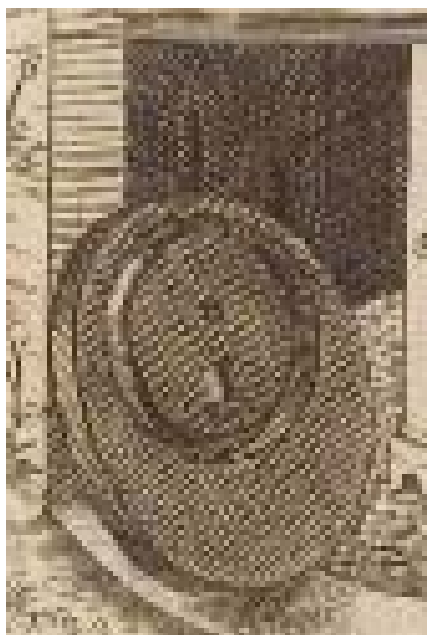
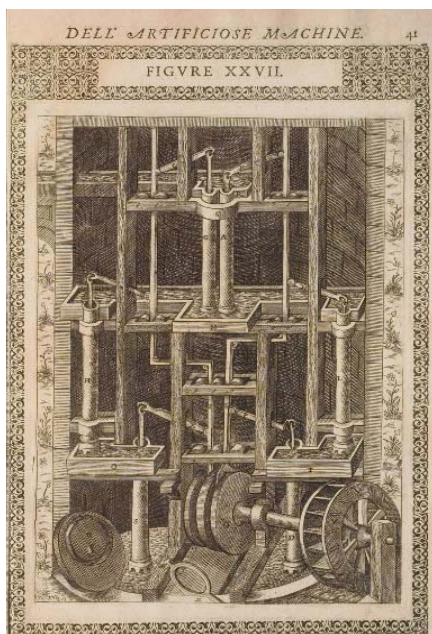


Figura 4. A sinistra: “Machina per far montare l’acqua” di Agostino Ramelli, 1588 [6]; a destra: Particolare della camma a comando positivo. Il piolo del cedente si impegna nei solchi di due camme affacciate.



Figura 5. Gualchiera, dal “Theatrum Machinarum Novum” di Andreas Böckler, 1661 [7].

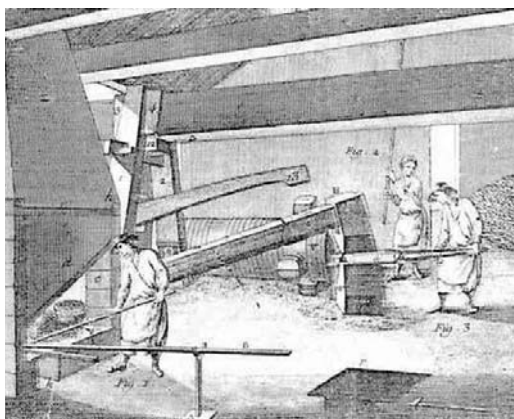


Figura 6. Fucinatura dei lingotti in un'officina del XVIII secolo. Enciclopédie Diderot, moulinaffer.free.fr.

3. CARILLON, SCATOLE MUSICALI, ORGANI MECCANICI

Carillon di campane

In Occidente, il più antico carillon di campane di cui si ha notizia pare sia quello della Cattedrale di Strasburgo del 1352; un altro importante carillon è quello costruito nel 1554 da P. van den Ghey: le parti sopravvissute di questo carillon sono ora al Rijkmuseum di Amsterdam.

La massima diffusione dei carillon si ebbe però dal '600 al '900, soprattutto in Belgio, in Olanda e nel Nord della Francia, ma anche in molte altre parti d'Europa [10].

I tamburi dei carillon, come quelli delle scatole musicali e degli organi meccanici, si possono considerare derivati direttamente dai primitivi alberi a camme. Questi tamburi portano in-

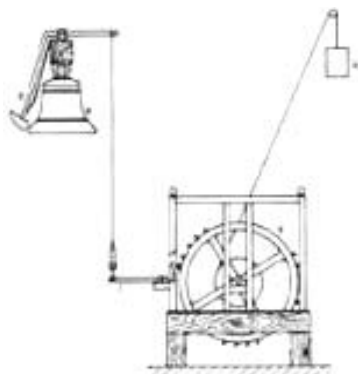


Figura 7. A sinistra: Schema dell'azionamento di una campana di carillon; a destra: Tamburo del grande carillon di Ghent, 1659 [10].

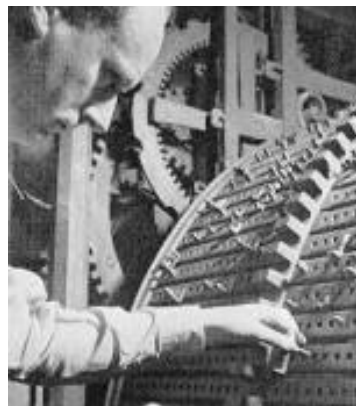
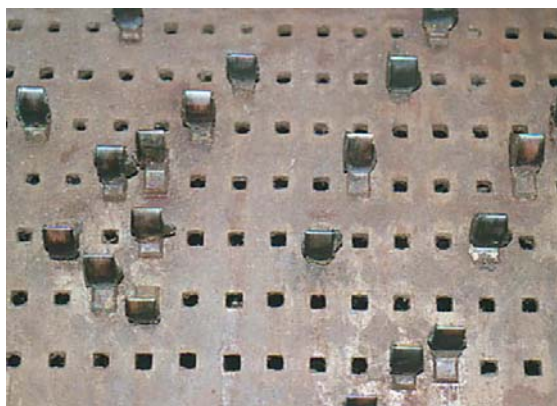


Figura 8. Particolare di tamburo di carillon e inserimento di un nuovo motivo (www.msu.edu).

fatti delle file di pioli disposti su una circonferenza: ogni fila aziona il meccanismo del batacchio di una campana, v. Fig. 7. Anche in questo caso le camme, cioè i pioli, devono imprimere al cedente solo un determinato spostamento, e non una particolare legge di moto.

I carillon possono essere programmabili, è cioè possibile cambiare il motivo musicale: in questo caso, il tamburo porta semplicemente dei fori, entro i quali i pioli che costituiscono le camme possono essere disposti in vario modo, v. Fig. 8.

Scatole musicali

Analoghi a quelli dei carillon sono i tamburi delle scatole musicali [11], che hanno avuto larga diffusione soprattutto dal Settecento in poi. Le scatole musicali sono peraltro sostanzialmente più semplici dei carillon di campane, perché i perni del tamburo agiscono di-



Figura 9. Scatola musicale.

rettamente sulle laminette sonore che, accordate sulle varie note, emettono i suoni richiesti, v. Fig. 9. Affinché sia possibile cambiare il motivo musicale, si deve poter sostituire il tamburo con un altro, diversamente programmato.

Organi meccanici

Gli organi meccanici sembra abbiano avuto illustri antenati nell'antichità: viene infatti citato un organo a vento di Erone, di epoca ellenistica, del quale sono state tentate anche moderne ricostruzioni grafiche in base alle descrizioni pervenute. Qui però considereremo solo gli organi meccanici diffusi e documentati in epoca moderna.

Come nei carillon di campane, anche negli organi meccanici [11] i tasti sono azionati da un tamburo, v. Fig. 10; qui, però, è possibile variare la durata delle note, per cui le sporgenze non sono dei semplici pioli, ma possono avere lunghezze differenti. Poiché l'organo richiede dei mantici per creare aria compressa da inviare nelle camme, sono stati realizzati anche organi idraulici, nei quali sono azionati idraulicamente sia il tamburo, sia i mantici.

Da rilevare che nel passato sono state costruite delle macchine musicali molto complesse, vere e proprie "orchestre meccaniche", in grado di produrre le voci di molti strumenti, sia a percussione sia a fiato e a corda. Veri precursori del grammofono e

del juke-box, queste macchine musicali venivano utilizzate sia in luoghi di intrattenimento pubblico, sia in case private.

4. AUTOMI

Un notevole precursore dei successivi automi meccanici realizzati in Europa nel Settecento è il dispositivo descritto da al-Jazari nel 1276 [12], che è contemporaneamente un automa e una scatola musicale programmabile, v. Fig. 11. In questo "battello da tavola", destinato ad allietare i banchetti, un albero a camme fa muovere quattro suonatori; cambiando le camme, si può cambiare il motivo suonato.



Figura 10. Organo meccanico [11].



Figura 11 – Battello di al-Jazari, 1206. La ruota idraulica trascina l'albero a camme, e queste fanno muovere i quattro suonatori visibili sulla sinistra del disegno.

Nel battello di al-Jazari le camme sono ancora del tipo primitivo. Come abbiamo visto, nei secoli successivi, qualche piccolo progresso era stato compiuto in alcuni settori, per esempio con l'impiego di camme di lunghezza variabile negli organi meccanici o con la camma a comando positivo del Ramelli, peraltro solo ideata. Nella seconda metà del Settecento, invece, le camme fecero un improvviso e straordinario salto di qualità, per opera di alcuni costruttori di automi, quali Vaucanson, Kintzing, i Jaquet-Droz, Maillardet.

Il grande successo degli automi va probabilmente inquadrato nella particolare temperie culturale dell'Epoca dei lumi: l'automa, infatti, fu considerato un simbolo della capacità dell'uomo di dominare la natura, di imitarla alla perfezione e addirittura – con qualche palese esagerazione – di diventarne egli stesso creatore.

Uno dei primi e più significativi esempi è l'automa musicale noto come “La joueuse de tympanon” [13] o suonatrice di dulcimero, realizzata nel 1784 da Kintzing per la regina Maria Antonietta. In questo automa, delle camme complesse, analoghe a quelle delle moderne macchine automatiche, imprimevano i movimenti al corpo e alle membra della piccola suonatrice, mentre il movimento di battuta dei martelletti è ottenuto per mezzo di pioli disposti su un tamburo. La scelta di suonare un semplice strumento a percussione permette di evitare l'esigenza di modificare la durata e l'intensità delle note, mentre per cambiare il motivo basta spostare assialmente il rullo, sul quale le camme corrispondenti ai diversi motivi sono disposte su circonferenze parallele.



Figura 12. “La joueuse de tympanon”, realizzata nel 1784 da Kintzing e Roentgen, in grado di suonare otto diversi motivi musicali. ©Musée des arts et Métiers, Parigi.

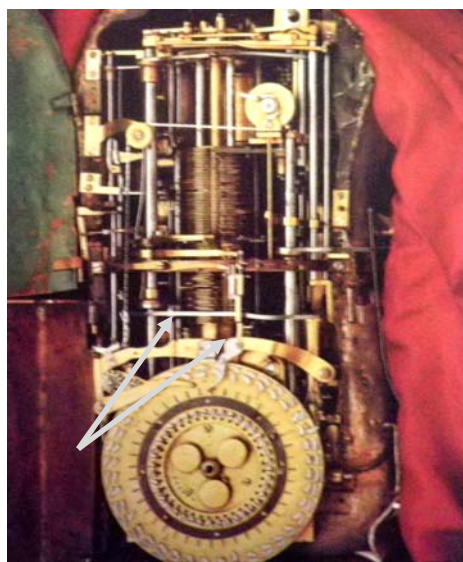


Figura 13. Lo scrivano dei Jaquet-Droz [15]. Nella figura di destra sono evidenziati i nottolini per la programmazione dei caratteri da scrivere.

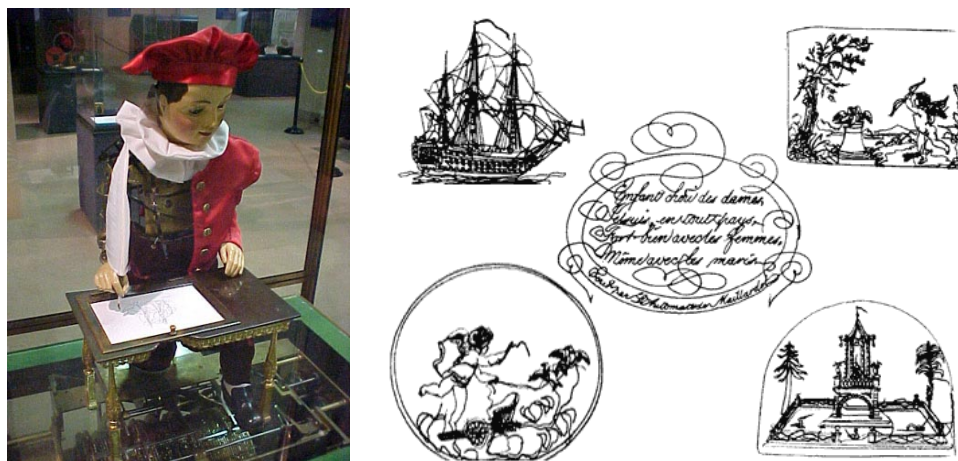


Figura 14. A sinistra: l'automa di Maillardet, ora al Franklin Institute di Philadelphia; a destra: la scritta e i disegni che l'automa è in grado di tracciare [14].

Gli automi meccanici più famosi sono senz'altro i tre Androidi [14], [15] costruiti tra il 1768 e il 1771 dall'orologiaio svizzero Pierre Jaquet-Droz, da suo figlio Henri-Louis e dal loro aiutante Jean-Frederic Leschot. Lo Scrivano, la Musicista e il Disegnatore sono automi davvero straordinari. Numerose camme imprimevano alle loro membra, al busto, alla testa, agli occhi movimenti ed atteggiamenti che imitano alla perfezione quelli umani; ed è proprio muovendo le braccia e le mani che scrivono, suonano, disegnano: come esseri umani, appunto.

La Musicista, per es., suona un organo, e le camme che la azionano permettono di modulare la durata del contatto fra le dita ed i tasti, cioè la durata della nota.

Il più complesso degli Androidi è però lo Scrivano, v. Fig. 13, dotato di un dispositivo per programmare il carattere da scrivere e, quindi, la successione dei caratteri e delle pause, per cui la frase da scrivere può essere cambiata senza cambiare le camme.

Molto complesso è anche l'automa costruito da Henri Maillardet [16], ex collaboratore dei Jaquet-Droz: un disegnatore-scrivano ora al Franklin Institute di Philadelphia, USA. Questo automa è in grado di disegnare e di scrivere, v. Fig. 14; si ritiene che esso possieda la più ampia memoria meccanica di tutte le macchine simili mai costruite.

Automi meccanici più o meno complessi sono stati costruiti anche in seguito, ma non hanno più raggiunto la perfezione e la magia di questi primi esempi.

5. CALCOLO ANALOGICO E SERVOMECCANISMI

Gli Androidi e gli altri Automi dimostrano che nella seconda metà del Settecento si possedevano conoscenze tecniche sufficienti per risolvere i problemi fondamentali della Cinematica geometrica. Le camme avrebbero dunque potuto essere impiegate già allora per risolvere problemi di calcolo analogico o per costruire macchine automatiche: questi problemi, evidentemente, non erano però ancora maturi, e non furono affrontati che molto più tardi.

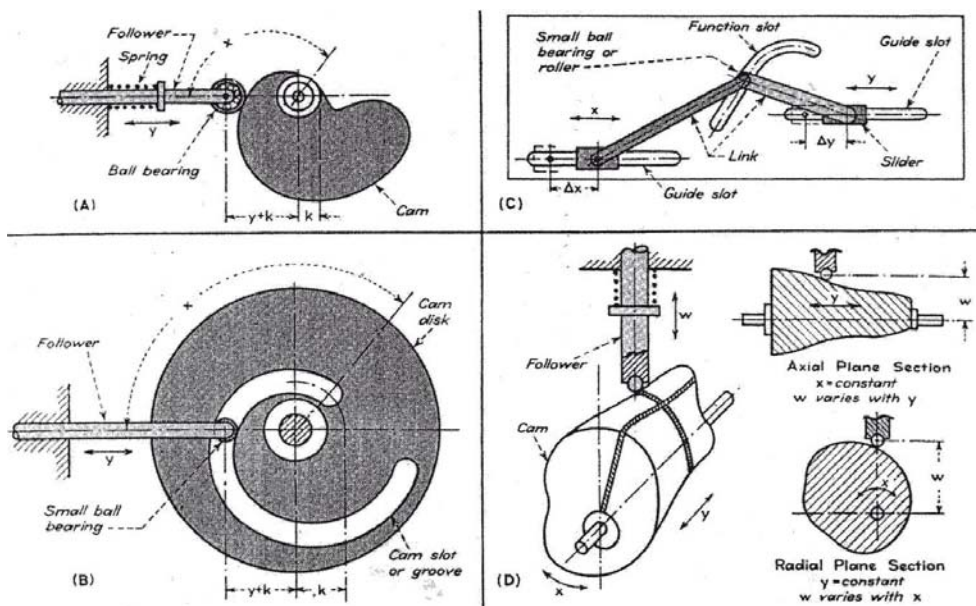


Figura 15. Camme per la generazione di funzioni [17].

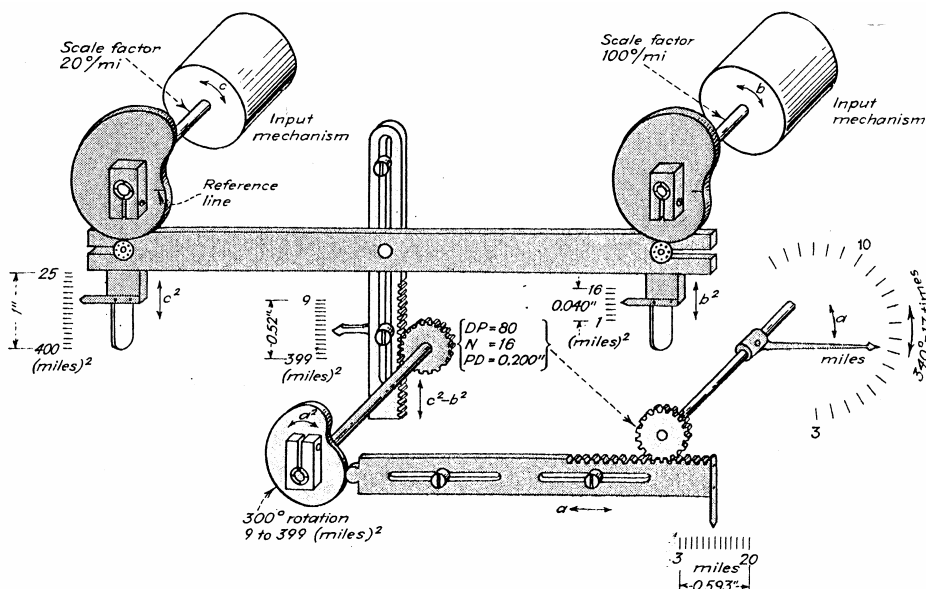


Figura 16. Schema di calcolatore meccanico analogico per risolvere la relazione

$$a = \sqrt{c^2 - b^2} \quad [17].$$

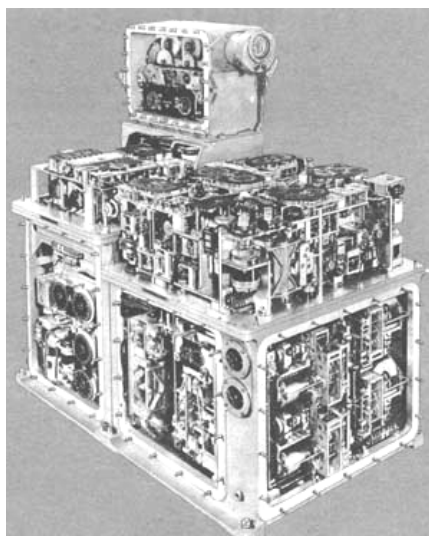


Figura 17. Il Computer MARK 1 della Ford Instrument Company, Long Island City, New York. Come parte del sistema MARK 37, era impiegato nei sistemi di puntamento dei cannoni calibro 5"/38, 5"/54, 6"/47, e 8"/55 per obiettivi di terra e di aria [18].

I primi impieghi pratici delle camme per i problemi sopra citati sembrano infatti risalire al periodo fra le due guerre.

Nella Fig. 15 sono riportati esempi di camme per la generazione di funzioni [17], utilizzate nei calcolatori analogici, nei servomeccanismi e nelle macchine automatiche, mentre nella Fig. 16 è rappresentato lo schema di un calcolatore analogico da utilizzare per ricavare la lunghezza di un cateto di un triangolo rettangolo, note quelle dell'ipotenusa e dell'altro cateto [17].

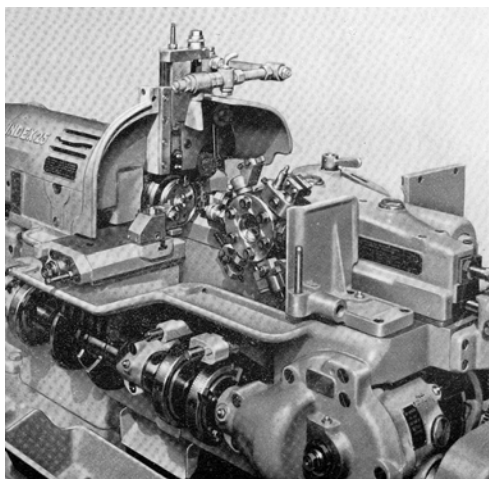


Figura 18. Torretta, slitte e camme di un tornio automatico Index [19]. Nei torni automatici le camme sono impiegate come generatori di funzioni.

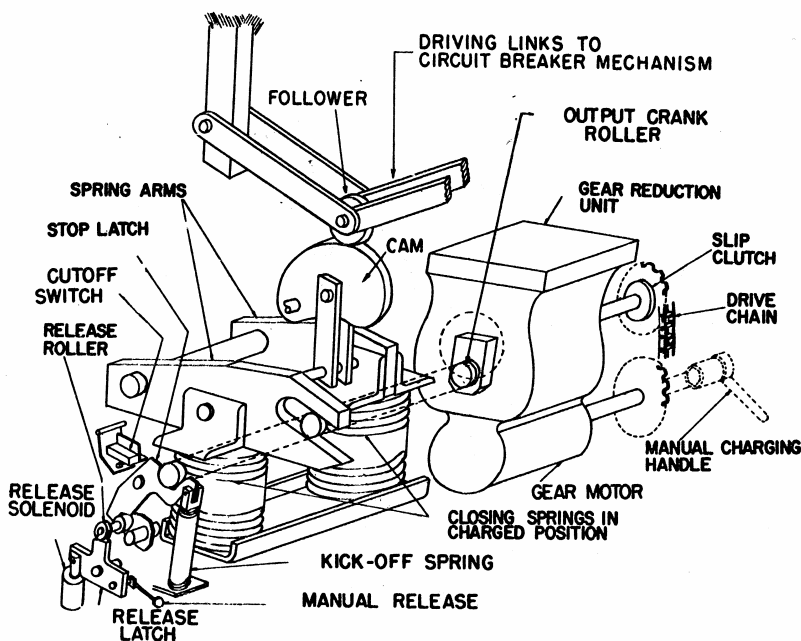


Figura 19. Schema di un interruttore elettrico di potenza della GE [20].

Dispositivi di questo tipo, basati peraltro sull'impiego sia di camme, sia di sistemi articolati, furono largamente impiegati nei servomeccanismi utilizzati per applicazioni militari. Un esempio è il MARK 1 [18], un computer analogico elettromeccanico usato dalla Marina Militare americana su portaerei, corazzate, incrociatori e cacciatorpediniere durante la seconda guerra mondiale, v. Fig. 17.

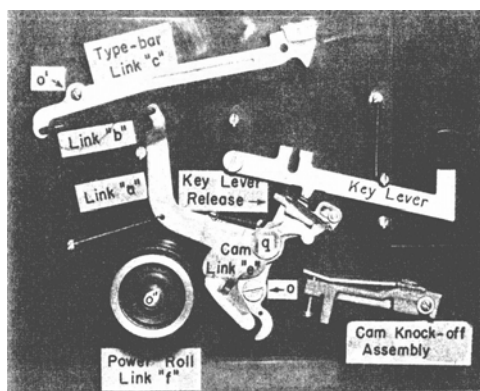


Figura 20. Meccanismo di macchina da scrivere elettrica IBM e suo schema cinematico [21].

6. ALTRE APPLICAZIONI

Nei primi tre quarti del secolo scorso le camme hanno avuto molte applicazioni: alcune sono ancora attuali e non vi è indizio che stiano per essere abbandonate, ma in altre si preferiscono ora soluzioni elettroniche, più efficaci e – soprattutto – molto più flessibili.

Citiamo, a titolo di esempio, i torni automatici [19], v. Fig. 18, gli interruttori elettrici di potenza [20], v. Fig. 19, e le macchine da scrivere [21], v. Fig. 20.

7. CONCLUSIONI

Le camme sono organi meccanici molto raffinati, che trovano oggi largo impiego in molte applicazioni, le più note e importanti delle quali sono senza dubbio le macchine confezionatrici automatiche e i motori a combustione interna.

Nel passato, esse sono risultate essenziali nello sviluppo di settori industriali di primaria importanza, quali la siderurgia, l'industria della carta, l'industria tessile e altri. Sono inoltre presenti in campi affascinanti come la musica meccanica e gli incantevoli automi del Settecento.

Lo studio delle loro applicazioni storiche costituisce dunque un viaggio estremamente interessante nella storia della tecnica e della cultura.

BIBLIOGRAFIA

- [1] <http://classes.bnf.fr/villard/feuille/index.htm> Folio 44.
- [2] Castagnaro G. (a cura di), 1990. *Contributi italiani alla diffusione della carta in Occidente tra XIV e XV secolo*. Fabriano, Pia Università dei cartai.
- [3] Taccola (Mariano di Jacopo detto il), 1433. *Liber tertius de ingeneis ac edifiitiis non usitatis*. Riproduzione in facsimile del ms. P.766 della B.N.C. di Firenze. A cura di J. Beck, Edizioni Il Polifilo, 1969.
- [4] Vannoccio Biringuccio, 1540. *De la Pyrotechnia*. Biblioteca Digitale, fermi.imss.fi.it, IMSS Museo Galilei, Firenze.
- [5] Giorgio Agricola, 1563. *De Re Metallica*, traduzione italiana di M. Florio, facsimile a cura di P. Macini e E. Mesini, Edizioni ANIM, 1994.
- [6] Agostino Ramelli, 1588. *Le diverse et artificiose machine del Capitano Agostino Ramelli dal ponte della Tresia ingegniero del Christianissimo Re di Francia et di pollonia*. Parigi.
- [7] Georg Andreas Böckler, 1661. *Theatrum Machinarum Novum*. Verlegung Paulus Fürsten, Gedruckt bey Christoff Gerhard, Nürnberg.
- [8] Museo del Patrimonio Industriale, Bologna, museopatbiblioteca@comune.bologna.it
- [9] Museo provinciale degli usi e costumi, Brunico, www.museo-etnografico.it
- [10] Huybens G., 1994. *Carillons et Tours de Belgique*. Ghent, Belgium, Ludion Editions, ISBN 90-5544-019-1.
- [11] Bowers Q. D., 1972. *Encyclopedia of Automatic Musical Instruments*. The Vestal Press, USA, ISBN 0 911572 65 1.
- [12] al-Jazari, 1206. *The Book of Knowledge of Ingenious Mechanical Devices*. Facsimile, 1990, Ankara, Publication of the Ministry of Culture.
- [13] <http://youtube.com/watch?v=75CXFwgsIsY> La Joueuse de Tympanon.

- [14] Maingot E., 1959. *Les Automates*. Parigi, Hachette.
- [15] Carrera R., Loiseau D., 1980. *Androidi*. F. M. Ricci, ISBN 88 216 0317 2.
- [16] Chapuis A., Gélis E., 1928. *Le monde des automates. Etude historique et technique*. Chez les Auteurs, Paris.
- [17] Reid R. R., Du Ray E. Stromback, 1949. "Mechanical Computing Mechanisms". *Product Engineering*, I: 131-135, Aug. 1949; II: 119-123, Sept. 1949; III: 126-130, Oct. 1949; IV: 121-124, Nov. 1949.
- [18] Ben Clymer A., 1993. "The Mechanical Analog Computers of Hannibal Ford and William Newell". *IEEE Annals of the History of Computing*, **15**(2), pp. 19-34.
- [19] Rossi M., 1980. *Macchine utensili*. Hoepli, Milano, ISBN 88 203 1162 3.
- [20] Barkan P., 1958. "High-speed spring-actuated cams". Transactions of the Fifth Conference on Mechanisms, Cleveland, The Penton Publishing Co. (64-76).
- [21] Droke R. L., 1948. "Pivoted-follower cam systems". Transactions of the Fifth Conference on Mechanisms, Cleveland, The Penton Publishing Co. (42-49).

A STUDY ON THE REHABILITATION HAND EXOSKELETONS

Mohammad Mozaffari Fomashi

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: mohammad.mozaffari2@unibo.it*

Marco Troncosi

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: marco.troncosi@unibo.it*

Vincenzo Parenti Castelli

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: vincenzo.parenti@unibo.it*

Abstract. *Within the frame of BRAVO (a project that aims at the development of an advanced robotic system for the rehabilitation of post-stroke patients with diseases at the upper limb), the authors are directly involved in the design of a hand exoskeleton (or active hand orthosis) intended to restore the patient's ability to grasp objects as a fundamental step for the rehabilitation of functional autonomy in the everyday living. The analysis of the state-of-art is the first step of our activity, preceding the design phase. This paper proposes a systematic classification of the reviewed systems based on two main key issues, namely the degrees of freedom of a finger exoskeleton and the number of contact points with the human finger phalanges. This classification is helpful to understand the reason of proposing certain solutions for the different applications and the advantages and drawbacks of the different designs proposed in the literature.*

Keywords: *Rehabilitation Robotics, Hand Exoskeleton, Finger Classification*

1. INTRODUCTION

The hand is an organ of grasp as well as of sensation, fine discrimination and exquisite dexterity. Unfortunately hand injuries are very common and the number of people suffering from that is increasing nowadays [1]. The injuries include traumatic injuries, problems related to the ageing (e.g. arthritic conditions), congenital deformities and the problems of a weak control of movement and force due to neurological diseases. The corresponding malfunctioning of the hand results in limitation in activities of daily living. The importance of having a normally functioning hand for an independent and active life needs no emphasis. At the rehabilitation phase, the recovery of hand functioning is generally assisted by physical therapists that make the patients perform some exercises. These exercises are mainly concentrated on basic gross mobility skills such as moving specific joints and

strengthening specific muscles. Further exercises, more complicated, can possibly be executed in order to allow the patient to recover the ability in activities of daily living (ADLs). A few months are required quite often to achieve acceptable improvement in finger movement and autonomous control. So the rehabilitation procedure is generally time consuming and with high cost. A way to solve part of this problem is doing rehabilitation and therapy exercises at home. Although it could be a positive solution, it should be noted that no monitoring and no systematic control on the therapy practices is a problem. Lack of motivation of patients to perform exercises at home can be also another problem.

With focus on the literature of the last decades, there has been a growing interest and research activities to explore the use of robotic systems in the area of rehabilitation purposes to solve (in part) the mentioned problems. In particular, a number of hand exoskeletons (HEs), also known as active hand orthoses, have been proposed as robotic aids for substituting or assisting the therapists in the rehabilitation process [1-15]. The main issue that must be taken into account in the use of hand rehabilitation robotic systems is the factor "Assistance-as-needed". In a few words, the exoskeleton should provide as much force/movement assistance as needed to accomplish the task, whereas it should enable the patient who does not need power assistance to move autonomously. In this latter case, the machine should not contrast the finger "free" movement.

According to various types of hand problems, different rehabilitation protocols and HEs have been proposed. Obviously, also the design of the robotic systems strictly depends on the functions to be accomplished and the rehabilitation procedure to be implemented. Considering the implications on the exoskeleton design issues, three aspects of the rehabilitation protocols could be pointed out that can lead to completely different solutions:

- *global complexity of exercises*: basic movements of the finger joints vs. execution of ADLs. The basic movements of the finger joints consist of simple exercises intended to improve the finger mobility and strengthen the muscles. Performing manipulation tasks related to ADLs (e.g. grasping a bottle and pouring liquid into a glass) involve cognitive aspects as well and is intended to restore also the ability to control autonomously one's hand movements;
- *complexity of the hand movements*: gross motion of all the fingers (e.g. for a cylindrical grasp) vs. single finger dexterity. Rehabilitation of power grasping (in which all fingers are actuated simultaneously) is one of the main features that must be rehabilitated, whereas the rehabilitation of the finger dexterity is slightly harder objective, requiring the capacity of the brain to perform the so-called "fractionation" that is the ability to move the fingers independently (and this could be a problem for patients with brain injuries);
- *rehabilitation environment*: real environment vs. virtual reality. When performing exercises of ADLs, rehabilitating a patient in a virtual environment is closer to what he/she will do out of the rehabilitation procedure. On the other side, the great advantages of virtual reality is that a simulator software can create many different environments and quickly switch between them without any set-up time. In addition, it makes motivation to the patient to do the exercises because it looks like playing a game.

It is worth mentioning that with respect to other kinds of robotic hands, a HE is directly attached to the human hand so that the movements of the two "systems" are coupled while – from the dual point of view – forces/moments are exchanged. In the design of such machines, and starting from the specific tasks that must be performed, a number of critical

issues are therefore to be considered, e.g. the control of the transmitted forces is mandatory for safety reasons, the motion of the HE links must be consistent with that of the human fingers and so on.

As for the design of every new technological products, the analysis of the state-of-art may help in defining some key issues and related solutions, by observing the existent products/prototypes that have been proposed so far. In a thorough review of the literature, a systematic approach would greatly help to understand the close relation among the different key factors and aspects of designing. In this perspective, the present paper proposes a systematic classification, focusing on the kinematics of the system, based on two main key issues: number of degrees of freedom (DOFs) and number of contact points (or connections) between human finger phalanges and the HE. This classification is helpful to understand both the reason of proposing certain solutions for the different applications and the advantages and drawbacks of the different designs proposed in the literature. This activity was carried out in order to obtain useful guidelines for the design of new HEs.

2. THE REHABILITATION HES

From a kinematic point of view and according to a popular model, the human hand has 20 degrees of freedom (DOFs) [16]: indeed each finger has 3 joints and 4 DOFs each. In particular, for the index, middle, ring and little fingers the joints are called (starting from the distal end) distal interphalangeal (DIP), proximal interphalangeal (PIP), and metacarpophalangeal (MCP) joints. The DIP and PIP joints have flexion/extension DOF, while the MCP joint has both flexion/extension and abduction/adduction DOFs. For the thumb the terminology is slightly different: from the distal end there are the interphalangeal (IP), the metacarpophalangeal (MCP) and the Carpometacarpal (CMC) joints. The IP and MCP joints have flexion/extension DOF, whereas CMC joint has both flexion/extension and abduction/adduction DOFs.

Depending on the applications, different HEs present different kinematic characteristics, in terms of both actuated DOFs and transmission chain layout. Generally speaking, HEs control the movement of fingers by means of a series of links which could be placed above the fingers, beside the fingers or inside the palm. Each solution contains some advantages and drawbacks according to their placement. First critical problem that must be solved for those devices that are placed above the fingers is the mechanical interference between the links and the anatomical parts. A number of devices use different kind of Remote Centre of Motion (RCM) mechanisms, such as gear-rack and slider mechanism, to make the mechanism centre of motion coincide with the human joint axis while avoiding mechanical interference. Other devices that are placed above the fingers do not exploit any RCM mechanisms and use simpler solutions (such as 4-bar mechanisms) to transmit motion and avoid interference. The exoskeletons that are placed beside the fingers do not need any RCM mechanisms [17], because their center of rotation is coincident with the center of rotation of the human finger. It looks great benefit, but this solution needs space beside the fingers that are not always available (e.g. between index and middle, middle and ring fingers). The exoskeletons that are placed inside the palm do not need any RCM mechanisms [18] as well. They could be very simple and light weight. The main disadvantage of these HEs is that their placement will decrease the workspace of the exoskeleton and makes complete grasping impossible.

Apart from the kinematic scheme of the device, another discriminating characteristic of different HE solutions is the choice of actuators and transmission systems. The

rehabilitation HEs typically use pneumatic actuators or electric actuators (whereas no hydraulic systems were found in the literature). The HEs that use pneumatic actuators could be more lightweight than those using electric actuators. But, in general, they need a big external compressor which is a major drawback. In order to keep the weight of the exoskeletons limited, the actuator parts are usually far from the exoskeleton (e.g. placed on a frame fixed to the user's forearm). The power is transmitted to the exoskeleton attached to the fingers by means of different transmission systems such as wire-driven mechanisms or linkage mechanisms. It is worth mentioning that in some devices the use of very lightweight electric motors (ultrasonic motors) makes it possible to place the actuator directly above the fingers [19].

The control of a HE is generally in charge of the wearer, who must generate proper input signals (basically depending on the application for which the system is designed) to control the HE finger functions, e.g. EMG-signals gained from human muscles [2, 8, 20, 21] or the movements measured on the other hand (using the so-called self motion control, [6]) or the force exerted on the HE itself [22]. In addition to the input command sensors, exoskeletons need some position sensors and force sensors to implement proper motion control strategies (typically position control or force control or a combination of both). For position controlling, most exoskeletons use incremental encoders integrated in the electric motors and the joint angles are calculated by means of the kinematic scheme of the mechanism. Some exoskeletons use sensors (e.g. potentiometers, Hall effect sensors, absolute encoders) directly placed on joints to measure the relative angular position of the connected links. Also the force sensors can be placed in different parts of the exoskeleton, depending on their characteristics and working principle. In some devices, they are directly placed between the HE and the human fingers to measure the contact force. Unfortunately, in this arrangement the sensors do not distinguish between forces exerted by the user and external forces. Thus, during contact with the environment, it becomes impossible to recognize the user's intention. Additional force sensors at every possible contact surface could detect external forces, but would trigger losing direct contact. In other cases, strain gages are placed on some components (e.g. to measure the tension of cables in wire-driven mechanisms) and the grasping force is calculated based on the retrieved data. Finally in some systems, the force exerted to the human fingers is computed and controlled by measuring and controlling the current draining of the electric motors. Unfortunately, modeling or measuring the force lost by friction (that contributes to the amount of the motor driving torque) is a very difficult task, so that the systems that exploit this strategy are generally not very accurate in force control.

In the following, we briefly describe the architecture, the systems' kinematics and the control schemes of 4 different rehabilitation HEs that are considered particularly significant (in their different design solutions) among the systems that were found in the scientific literature [1-15, 20-22].

Berlin University Hand

The device was designed with focus on support of the rehabilitation process after hand injuries or stroke [2, 21, 22]. The control method exploits Electromyography (EMG) sensors located on palmer and dorsal side of the forearm.

This HE control the movement of a hand in 20 DOFs. Each finger is supported in 4 DOFs: flexion/extension in MCP, PIP and DIP joints; abduction/ adduction in MCP joint.

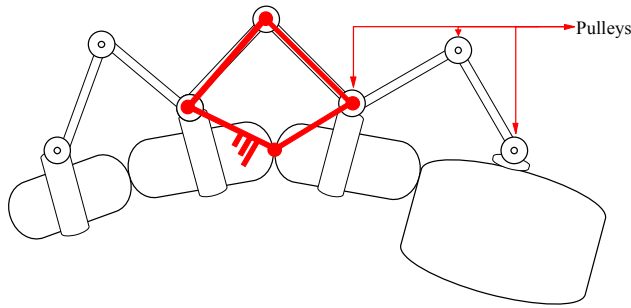


Figure 1. Scheme of one finger exoskeleton, based on a four-bar mechanism.

The thumb can also be moved in 4 DOFs. The CMC joint is supported in flexion/extension and in abduction/adduction movement, whereas the MCP and IP joints are supported in their flexion movement. The exoskeleton moves finger by means of 3 four-bar mechanisms. The joints are driven through wire-driven mechanism. One actuator unit consists of 4 DC motors for actuating 1 finger, so that five actuator units are required for a hand.

Each four-bar mechanism contains 2 external links and 2 human finger links. Further, it contains 3 external joints (2 Free joints and 1 actuated pulley) and 1 human joint (DIP or PIP or MCP joints), as represented in Fig 1. Each four-bar mechanism is freedom 1-DOF system which could be controlled via an actuator directly or indirectly such as through wire-driven mechanism, connected linkage mechanism etc.

The system is equipped with 5 different kinds of sensors: 2 angular measurement sensors, 1 force sensor, 1 Electromyograph sensor and 1 electrical current sensor. Electromyography electrodes measure the electronic activity at the skin up to 16 muscles, which is used as input for controlling the machine. To provide angular position measurements, the angle between 2 external links in each four-bar mechanism is measured by Hall sensors which are evaluated by integrated circuits. On the other side, optical quadrature encoders are used to measure the position of the motor axes. Three force sensors are placed between the human finger attachments and the exoskeleton. Motors are controlled by PWM-controllers. The actual current of the motor deviate from the values set by the control system. By these current, the torques at the motor axes are computed as well. Together with this torque and mechanics of linkages, the force exerted to the phalanges could be evaluated.

This HE uses two different parallel sensors for position analysis and two different parallel sensors for force analysis. It is possible to detect mechanical failure by comparing values gained by each sensor and correct them. Detailed information about the control procedures are explained in [22].

HWARD

The device was designed to support the rehabilitation process after hand injuries [3]. The exoskeleton is pneumatically actuated and supports repetitive power grasping and releasing movements of the human hand.

The device controls the movement of fingers in 2 DOFs and the wrist in 1 DOF. The combination of 4 fingers is controlled in 1 DOF for the flexion/extension about the MCP and PIP joints. The device is attached to the combination of 4 fingers along the dorsal side of the fingers by a link. This link is fastened to middle phalanges of 4 fingers by looping a soft strap around the 4 fingers. Movement of this link causes the movement of middle phalanges of combination of 4 fingers about the MCP and PIP joints. The movement of middle phalanges causes the movement of MCP and PIP joints because of internal DOFs among human joints. Distal phalanges and DIP joints of fingers are kept free. Similarly to the group of the 4 fingers, the thumb is actuated in 1 DOF for the flexion/extension about the CMC and MCP joints. The device is attached to the thumb along the dorsal side of it by a link. This link is fastened to the proximal phalanx of thumb by a soft strap. Movement of this link causes the movement of the thumb about the CMC and MCP joints. Distal phalanx and IP joint of the thumb are kept free.

Robot joint movement is achieved using a lever design. Each air cylinder and limb interface are mounted on opposite ends of the lever, with a revolute joint between them. Air pressure causes the movement of limb interface that causes the movement of levers connected to the human fingers. The system is actuated by 2 double-actuating air cylinders that can provide up to 122.8 [N] of force, but air pressure is regulated so that the air cylinders produce roughly 4-15 [N], the estimated level necessary to assist movement.

The system is equipped with 2 kinds of force sensors (pneumatic solenoid valve and micro structure pressure sensors) and 1 kind of position sensor (rotary potentiometers). The pressure of air injected into the cylinders is controlled by means of pneumatic solenoid valves and micro structure pressure sensors that are mounted on both sides of each air cylinder. The forces applied by the robot can be computed from data gained from these pressure sensors. In addition to pressure control, inline flow control valve adjust the rate of air injected to the cylinders. Rotary potentiometers measure the fingers, wrist and thumb joint angles.

The control procedure is only based on position controlling and the main focus is on rehabilitation of power grasping.

Pittsburgh University Hand

The device was designed with focus on the rehabilitation of natural pinching [4, 20]. The control method exploits EMG signals to control the movement of the index finger. The exoskeleton controls the movement of the index finger in 2 DOFs. The flexion/extension movement of PIP and DIP are coupled and actuated via a pneumatic piston and the flexion/extension movement of MCP joint is actuated via another pneumatic piston. The index finger has another passive DOF in the abduction/adduction movement. Pinching motion is performed via flexion of index finger against a fixed thumb. The frame of the exoskeleton consists of an aluminum anchoring plate mounted to the back of the hand and three aluminum bands, one for each phalanx. The flexion of the PIP and DIP joints is produced by a steel cable running along the front of each finger band and through to the backside of the hand. This cable is pulled by a pneumatic cylinder acting in compression. The extension movement of the PIP and DIP joints are produced by 2 springs as shown in Fig. XX. The MCP flexion is achieved by a linkage mechanism: a floating link is mounted between the finger band closest to the base plate and a second pneumatic actuator. When this pneumatic piston pushes the link mechanism forward (distal), the MCP joint is flexed. The extension movement of the MCP joint is produced by a spring (Fig. 2). To achieve

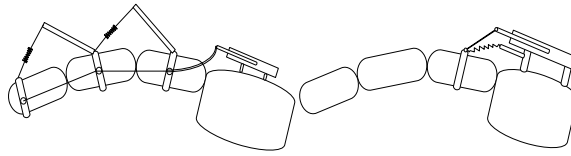


Figure 2. The schematic model of HE.

smooth repeatable motion and the passive abduction/adduction motion, the flexible coupling added between the base-plate and first finger band is made of a canvas-like cloth material.

The exoskeleton movements are controlled by the users biceps EMG signals. The pistons are equipped with variable pressure pneumatic valves to control the maximum forces exerted to the finger phalanges. The system uses two main algorithms to control the output of pneumatic cylinders: binary control algorithm and variable control algorithm. The algorithm is called a binary control when the output to the pneumatic valves are set either off (0V) or on (10V). The output binary value was determined by the EMG signal. When the EMG signal was above a specified EMG threshold value, the output was “on”, and when it was below, the output was “off”. The force exerted by the exoskeleton is fixed before the rehabilitation procedure with a calibration trial. The algorithm is called a variable control when the output to the pneumatic valves are set in 3 situations: off, minimum pressure level (20 psi), maximum pressure level (120 psi). The EMG signal is filtered in this algorithm. If it is at 15% of the maximum muscle contraction level, minimum pressure level (20 psi) are set and if it is at 70% of the maximum muscle contraction level, maximum pressure level (120 psi) are set on output of pneumatic valves.

The system is mainly focus on repeatedly movement of index finger in pinching. The system does not equipped with any position sensors. So, there is not any control on position of the finger joints.

HANDEXOS

The device was designed with the purpose of supporting the rehabilitation process after stroke [5]. The exoskeleton uses an under-actuated mechanism that enables the hand to passively adapt each finger to the generic shape of the grasped object (self-adaptation).

The 5 fingers HE controls the movement of a hand in 5 DOFs. Each finger is supported in 1 DOF: the flexion/extension of DIP, PIP and MCP joints are actuated by means of a single DC motor and the abduction/adduction movement of MCP is supported passively.

A crucial aspect must be explained in the architecture of these kind of exoskeleton, which is placed over an operator finger. To mimic the motion of the operator, the rotation centers of the exoskeleton machine should coincide with the rotation centers of the operator fingers to avoid mechanical interference. Through avoiding this interference, many kinds of Remote Center of Motions mechanisms (RCMs) can be used. RCMs are mechanisms that are able to implement the rotation of a body around a fixed axis that is remotely located from the structure of the joint. This HE exploits 2 different strategies to make the exoskeleton joint axes coincide with those of the human finger. For DIP and PIP joints, 4 pulleys, 2 for each joints (one for actuating the flexion and one for extension), are placed on both sides of HANDEXOS finger at the level of the wearer's finger joint. For the MCP joint, 2 pulleys are placed above the human MCP joint to actuate flexion/extension

movement of finger. These pulleys are connected to the first phalanx by a passive prismatic joint. By using this slider-crack mechanism, the length of the link could be adjusted passively to avoid interference between exoskeleton and human finger during the flexion movement of MCP joint.

The actuator unit consists of 1 DC motor that is used to extend the coupled DIP, PIP and MCP joints. The extension movement of each finger is actuated by a cable running across idle pulley placed in each finger joints and fixed to the distal phalanx through a cable stop. The cable is pulled through a linear slider by a DC motor placed extrinsically. The flexion movement of the finger is passively obtained by means of a set of three (one for each joint) antagonist cables running across the pulleys placed on the other side of the finger, connected to three extrinsic linear compression springs whose elastic torques cause the finger to flex.

The main advantages of the under-actuated mechanism that is used in this HE is its capability of self-adaptation. Indeed, the finger is enabled to passively adapt each finger to the shape of the grasped object (it is worth recalling that the flexion movement is provided by the elastic force of springs).

3. CLASSIFICATION

According to different applications and various kind of rehabilitation protocols, different rehabilitation HEs have been proposed in the literature. As it can be concluded by analyzing the four rehabilitation HEs mentioned in Section 2, one can find HEs that are completely different in their kinematic schemes, in particular with reference to the number of DOFs (actuated or not) and number and type of connections with the human fingers. In this perspective, some main questions can arise: which structure is suitable for a particular purpose?, how many DOFs or contact points are suitable for a particular HE? These questions are still widely open, in part because a systematic review of the different rehabilitation HEs is totally lacking. Thus, for better understanding the proposed design solutions, a systematic classification of all HEs could be helpful.

Here, one possible kind of classification of one exoskeleton finger (considered as the basic unit to analyze) is introduced based on two principles, i.e. number of DOFs and number of contact points (CPs) with the human finger, selected as discriminating characteristics for distinguishing a number of proposed solutions. Indeed these two aspects mainly determine the kinematic scheme of the device and its architecture, and are strictly connected to the system targeted application. This classification is helpful to understand the reason of proposing different solutions in different rehabilitation HEs and the advantages and drawbacks of a given solution.

Number of DOFs

Firstly it should be noted that the abduction/adduction movement of MCP joints is not considered in the analysis of the finger kinematics because the most interesting differences among the HEs regard the flexion/extension movements and most HEs do not actuate the abduction/adduction movement because in the rehabilitation process the efforts are concentrated in the flexion/extension. With this restriction, a human finger has 3 DOFs, so that a HE can provide a human finger with full mobility and versatility of movements only if it has 3 actuated DOFs. On the other side, it should be taken into account that the resulting device could be very heavy (basically due to the weight of the actuators) and the control procedure very complex. Thus alternative solutions can be obtained by accepting a

worsening of the system versatility and supporting the finger in less DOFs (1 or 2), thus making the global weight and the complexity of the system be limited. In this case, it is necessary to couple the movement of different joints or use under-actuated mechanisms and it should be noted that this choice can complicate the design of the transmission chain.

Based on these considerations we can identify three different groups of rehabilitation HEs, namely with 1, 2 or 3 DOFs per finger (flexion/extension movement of DIP, PIP or MCP joints), and compare the solutions proposed in the literature. With reference to the hands described in Section 2, the Berlin University Hand is extremely flexible as it controls 20 DOFs of the hand motion (4 DOFs per finger), but it is very complex and heavy. In our analysis, the finger of this HE is considered as having 3 DOFs. On the opposite side of the Berlin University Hand, the HWARD controls the movement of the combination of four fingers in 1 DOF. The Pittsburgh University Hand controls the movement of the index finger in 2 DOFs. The flexion/extension movement of PIP and DIP are coupled and actuated by 1 actuator and the flexion/extension movement of MCP joint is actuated by another actuator. Finally, the HANDEXOS uses an under-actuated mechanism to actuate each finger in 1 DOF: the flexion/extension of DIP, PIP and MCP joints are in fact actuated by means of a single actuator but there is no rigid coupling among the joints.

Number of CPs

The rehabilitation HEs needs to be connected to the human finger phalanges to control the movement of fingers and exert force on them (Fig. 3). The exoskeleton is attached to the finger by different methods. Usually it is attached by looping a strap or rigid ring around the finger phalanges [1-9, 13, 15] or using a cap fixed to the human fingertip [12]. In some HEs [10, 14], the user wears a glove and the rehabilitation HE is connected to the glove. In both cases, the exoskeleton and the finger phalanges do not move relatively one from the other. In a particular case [11], the device is attached to the finger by means of two rollers (above and beneath each phalange) that can move relatively to the finger phalanges. By using this relative movement, the exoskeleton does not need any RCM to make the mechanism centre of motion coincide with the human joint axis. Based on the number of CPs, we can identified three different groups: 1, 2 or 3 CPs per finger.

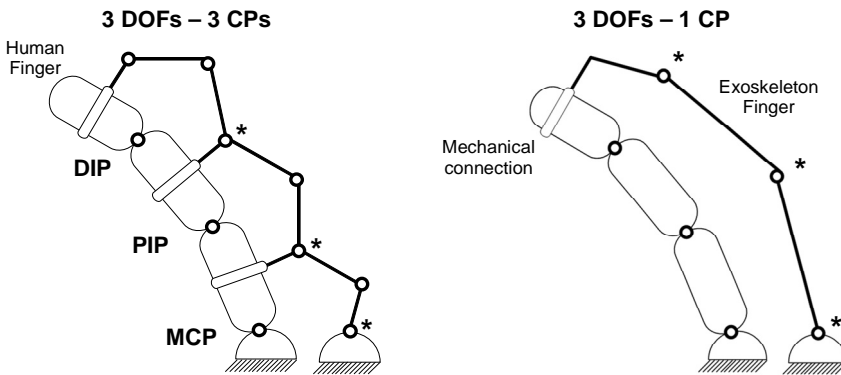


Figure 3. Schematic models of exoskeleton fingers connected to a human finger. The star symbols represent the presence of an actuator.

Number of CPs could be equal or greater than the number of DOFs in rehabilitation HEs. It should be taken into account that there is a close relation between the two different classification principles and selecting a certain combination of DOFs and CPs should be carefully based on a number of design criteria required by the specific application of the resulting HE. For instance, in Berlin University hand, the device is attached to the human finger phalanges in three points. Since the device controls 3 DOFs of the finger, the forces exerted on all the three phalanges are controllable. Also in the Pittsburgh University Hand, the exoskeleton finger is attached to the human finger phalanges in three points, but the device controls 2 DOFs of a finger only (being the DIP and PIP joints coupled). So, in this case, the force exerted to the distal and middle phalanges are not controllable separately.

According to these two classification principles, all rehabilitations HEs can be systematically arranged as shown in Tab. 1.

4. DISCUSSION AND CONCLUSION

The literature proposes a large number of HEs that generally present some common characteristics and many special peculiarities concerning their mechanics, electronics (control) and working principle. One of the main aspects that must be taken into account in designing new rehabilitation HEs – probably the first choice to be done – is relative to the kinematics of the system. Indeed, the designer must choose:

- the number of DOFs of the mechanism which guides the functions of a single finger (recalling that the human finger has 3 DOFs related to the phalanges' flexion/extension and 1 DOF for the MCP adduction/abduction);
- which joint movements must possibly be coupled (e.g. for a finger having 2 DOFs for flexion/extension, the coupling should occur between the DIP and PIP joints or between the PIP and MCP joints?);
- whether providing a number of actuators to control all the DOFs of the finger exoskeleton or designing an under-actuated mechanism (i.e. one or more DOFs are not directly controlled);
- how many CPs are needed to control the movement of the finger exoskeleton;
- how many finger mechanisms will form the overall HEs, and how many of them must be controlled independently.

Table 1. Classification of rehabilitation HEs .

	1 CP	2 CPs	3 CPs
1 DOF	- HWARD [3] - Milan University Hand [8]		- Harbin Institute of Technology Hand [1] - HANDEXOS [5]
2 DOFs		- Gifu University Hand [10] - Gentle/G Hand [7]	- University of Tokyo Hand [6] - Pittsburgh University Hand [4]
3 DOFs			- Berlin University Hand [2] - AFX Hand [11] - Beihang University Hand [9]

The different HEs proposed in the literature use different number of DOFs and CPs according to their applications, design constraints or characteristics. There is a device that controls 20 DOFs of the hand (complete movement of all fingers to replicate the overall human hand movements) [2]. Other devices use different coupling mechanisms to couple the movement of human finger joints and other systems leave some joints free (i.e. uncontrolled). Coupling between the joints of a single finger is done by means of different mechanical solutions such as cables and pulleys or rigid linkage mechanisms. In addition to the mentioned coupling in a single finger, some HEs move the combination of some fingers together (for instance, the 2nd, 3rd, 4th and 5th fingers are controlled by means of one actuator) [3, 8]. It is worth noting that in some HEs special under-actuated mechanisms are proposed as simple solution to approximately control the finger movement [5].

The study of rehabilitation HEs having significantly different characteristics was the motivation to do a systematic classification. One possible kind of classification based on two main characteristics that distinguish one system from another is introduced: number of DOFs and number of CPs. Based on these two principles, the rehabilitation HEs were systematically analyzed and classified (as shown in Tab. 1). But some questions are still remaining:

- How this classification could be helpful to understand the reason of proposing particular solution in alternative rehabilitation HEs?
- What is the difference among the solution found in the cell of Tab. 1 with 3 DOFs, 1 CP and the other cell with 3 DOFs and 3 CPs?
- Why some combination of DOFs and CPs are empty (i.e. no devices were proposed in the literature)?

Each solution (combination of DOFs and CPs) strictly depends on the target application of HEs, e.g. the kind of rehabilitation protocols to be implemented and exercises that must be performed (Section 1).

The links between the proposed classification and existing rehabilitation systems (mainly focused on the HEs that are presented in Section 2) are here discussed. For instance, if the main aim of the device is the rehabilitation of the power grasping and the focus is on assisting repetitive grasping and releasing movements, the best suitable solution is likely given by the combination of 1 DOF and 1 CP. More DOFs and CPs could trigger complexity of the system while the main aim of the device is rehabilitating the basic cylindrical grasping and the consequent releasing movement. Additionally, since the power grasp can be thought of as a 1 DOF task, it looks logical to combine the movement of four fingers (index, middle, ring and little) by means of a link and actuated them simultaneously (as it has been proposed in both devices in this cell [3,8]).

On the other side, let us consider a system whose main goal is training a safe extension/flexion movement of a single finger with focus placed on basic movement and single finger dexterity. Based on this goal, a good candidate solution is the one that combines 1 DOF and 3CPs. This is because we require the training on basic movements (simple extension and flexion) of a single finger so that 1 DOF is enough for this application (a significant versatility to perform every possible trajectory of a finger is not required). Additionally, with 3 CPs, we can guide all the phalanges of a single finger to reach a target configuration, as required by the rehabilitation exercise. In this case, the global force exerted to the finger phalanges is controllable which is enough for the goal of

the system (i.e. the independent control of the forces exerted by the HE on the 3 CPs is not necessary).

Finally, let us consider another rehabilitation protocols which focuses on the rehabilitation of a single finger not only for basic movement of the finger joints but also for the execution of a number of ADLs. For this application, the full versatility of the finger is required to accomplish the different trajectories required by the given ADLs. Therefore the exoskeleton must control the finger phalanges in 3 DOFs and 3 CPs. In this case, the forces exerted to the finger phalanges are fully controllable.

By analyzing the existing devices with this classification, and considering the real systems versus all the theoretical possibilities, it seems that some combinations of DOFs and CPs are not feasible (particularly when the number of CPs is lower than the DOFs). For instance it is possible to control 3 DOFs of a human finger and attaching the HE to the human finger in 1 CP only. This combination makes it possible to fully control the position and orientation of the driven phalange (typically the human fingertip), but it is not logical in rehabilitation HEs, where for the safety of the injured hand (the wearer shows some diseases and is not able to autonomously control the hand) the motion of the human joints should be fully guided. On the other side, it is worth mentioning that this combination is quite popular in haptic device applications.

Finally, it should be taken into account that a given decision on each principle contains some advantages and drawbacks and also it can have effects on other aspects of designing. In a few words, there is a close relation among the different key factors, and selecting a given specific solution can solve a group of problems related to one aspect of designing, but it could trigger other kinds of problems or limitations in other aspects.

In conclusion, this paper presents a review of the rehabilitation hand exoskeletons that are presented in the scientific literature, analyzing them by a classification that considers their kinematic schemes. In particular the number of DOFs of a single exoskeleton finger and the connection points with the human finger are considered as fundamental issues strictly related to the specific application of the resulting orthosis.

Acknowledgements

The research is performed in the framework of the project titled BRAVO "Brain computer interfaces for Robotic enhanced Action in Visuo-motOr tasks". The support of the Italian Institute of Technology (Genova, Italy) is gratefully acknowledged.

REFERENCES

- [1] Y. Fu, P. Wang and S. Wang, 2008. "Development of a Multi-DOF Exoskeleton Based Machine for Injured Fingers". *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, September 22-26, Nice, France, pp. 1946–1951
- [2] A. Wege, G. Hommel, 2005. "Development and Control of a Hand Exoskeleton for Rehabilitation of Hand Injuries", *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Berlin, Germany, pp. 3461–3466
- [3] C.D. Takahashi, L. Der-Yeghian, V.H. Le and S.C. Cramer, 2005. "A Robotic Device for Hand Motor Therapy After Stroke", *Proceedings of the IEEE 9th International Conference on Rehabilitation Robotics*, June 28 - July 1, Chicago, IL (USA), pp. 17-20

- [4] L. Lucas, M. DiCicco, and Y. Matsuoka, 2004. "An EMG-Controlled Hand Exoskeleton for Natural Pinching", *Journal of Robotics and Mechatronics*, **16**(5), pp.1–7
- [5] A. Chiri, F. Giovacchini, N. Vitiello, E. Cattin, S. Roccella, F. Vecchi, M.C. Carrozza, 2009. "HANDEXOS: towards an exoskeleton device for the rehabilitation of the hand", *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 11-15, St. Louis, USA
- [6] H. Yamaura, K. Matsushita, R. Kato, and H. Yokoi, 2009. "Development of Hand Rehabilitation System for Paralysis Patient – Universal Design Using Wire-Driven Mechanism", *31st Annual International Conference of the IEEE EMBS*, September 2-6, Minneapolis, Minnesota (USA)
- [7] R.C.V. Loureiro and W. S. Harwin, 2007. "Reach & Grasp Therapy: Design and Control of a 9-DOF Robotic Neuro-rehabilitation System", *Proceedings of the IEEE 10th International Conference on Rehabilitation Robotics*, June 12-15, Noordwijk, The Netherlands, pp. 757-763
- [8] M. Mulas, M. Folgheraiter and G. Gini, 2005. "An EMG-controlled Exoskeleton for Hand Rehabilitation", *Proceedings of the IEEE 9th International Conference on Rehabilitation Robotics*, June 28 - July 1, Chicago, IL(USA), pp. 371-374
- [9] J. Wang, J. Li, Y. Zhang and S. Wang, 2009. "Design of an Exoskeleton for Index Finger Rehabilitation", *Proceedings of the 31st Annual International Conference of the IEEE EMBS*, September 2-6, Minneapolis, Minnesota (USA)
- [10] H. Kawasaki, S. Ito, Y. Ishigure, Y. Nishimoto, T. Aoki, T. Mouri, H. Sakaeda, and M. Abe, 2007. "Development of a Hand Motion Assist Robot for Rehabilitation Therapy by Patient Self-Motion Control", *Proceedings of the IEEE 10th International Conference on Rehabilitation Robotics*, June 12-15, Noordwijk, The Netherlands, pp. 234–240
- [11] T.T. Worsnopp, M.A. Peshkin, J.E. Colgate and D.G. Kamper, 2007. "An Actuated Finger Exoskeleton for Hand Rehabilitation Following Stroke", *Proceedings of the IEEE 10th International Conference on Rehabilitation Robotics*, June 12-15, Noordwijk, The Netherlands, pp. 896-901
- [12] I. Sarakoglou, N. G. Tsagarakis and D. G. Caldwell, 2004. "Occupational and Physical Therapy Using a Hand Exoskeleton Based Exerciser", *Proceedings of IEEE/RSJ International Conference on intelligent Robots and Systems*, September 28 - October 2, Sendai, Japan
- [13] Y. Fu, F. Zhang and S. Wang, 2004. "Structure Types Design and Genetic Algorithm Dimension Synthesis of a CPM Machine for Injured Fingers", *Proceedings of the IEEE International Conference on Robotics and Biomimetics*, August 22 - 26, Shenyang, China
- [14] Z. Sun, X. Miao and X. Li, 2009. "Design of a Bidirectional Force Feedback Dataglove Based on Pneumatic Artificial Muscles", *Proceedings of the IEEE International Conference on Mechatronics and Automation*, August 9 - 12, Changchun, China
- [15] I. H. Ertas, E. Hocaoglu, D. E. Barkana and V. Patoglu, 2009. "Finger Exoskeleton for Treatment of Tendon Injuries", *11th International Conference on Rehabilitation Robotics of the IEEE*, June 23-26, Kyoto International Conference Center, Japan
- [16] B. Buchholz and T.J. Armstrong, 1992. "A Kinematic Model Of The Human Hand To Evaluate Its Prehensile Capabilities", *Journal of Biomechanics*, **25**(2), pp. 149-162
- [17] Y. Hasegawa, Y. Mikami, K. Watanabe and Y. Sankai, 2008. "Five-Fingered Assistive Hand with Mechanical Compliance of Human Finger", *Proceedings of the IEEE International Conference on Robotics and Automation*, May 19-23, Pasadena, CA (USA)
- [18] M. Bouzit, G. Burdea, G. Popescu and R. Boian, 2002. "The Rutgers Master II—New Design Force-Feedback Glove", *IEEE/ASME Transactions On Mechatronics*, **7**(2), pp.256-263

- [19] B. H. Choi, H. R. Choi. 2000. "SKK Hand Master – Hand Exoskeleton Driven by Ultrasonic Motors", *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Takamatsu, Japan, October 30 - November 5, Vol. **2**, pp. 1131 – 1136
- [20] M. DiCicco, L. Lucas and Y. Matsuoka, 2004. "Comparison of Control Strategies for an EMG Controlled Orthotic Exoskeleton for the Hand", *Proceedings of the IEEE International Conference on Robotics & Automation*, April, New Orleans, LA (USA)
- [21] A. Wege and G. Hommel, 2007. "Electromyography Sensor Based Control for a Hand Exoskeleton", *Proceedings of the IEEE International Conference on Robotics and Biomimetics*, December 15 -18, Sanya, China
- [22] A. Wege, K. Kondak, G. Hommel, 2006. "Force Control Strategy for a Hand Exoskeleton Based on Sliding Mode Position Control ", *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 9 - 15, Beijing, China

ANALYSIS OF THE INTERNAL TORQUE AND BENDING MOMENT ACTING ON THE LEGS OF THE 3-UPU TRANSLATIONAL PARALLEL MANIPULATOR

Ahmed Hachem Chebbi

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: ahmed.chebbi@mail.ing.unibo.it

Vincenzo Parenti Castelli

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: vincenzo.parenticastelli@mail.ing.unibo.it

Abstract. *The 3-UPU Translational parallel manipulator, where U and P are for universal and prismatic pair respectively, is a very well known manipulator that has been a benchmark for many issues. Many studies have been reported in the literature on singularities, workspace and joint clearance influence on the platform accuracy of this manipulator. However, thorough studies on its stiffness are still lacking. The stiffness of this manipulator that mainly depend on torsion and flexion of the limbs(legs) connecting the base with the movable platform of the manipulator. Much work has still to be done on the variation of torsion and flexion in the workspace of the manipulator. This paper investigates the distribution of torsion and flexion of the legs in the workspace and a procedure is presented in order to find the best location of a given trajectory in the manipulator workspace that is characterized by the highest stiffness.*

Keywords: *torque, bending moment, torsion, flexion, parallel manipulator*

1. INTRODUCTION

Parallel manipulators (PMs) have focused a great attention in the last decades for their complementary characteristics with respect to the serial manipulators. Indeed, they are characterised by high rigidity, high payload to the manipulator weight ratio, high dynamic performance versus a limited workspace and a low manipulability. Six degrees of freedom (DOF) parallel manipulators have been extensively studied. Moreover, recently great attention has been devoted to 3-DOF PMs since many applications need only three DOF. This family of manipulators can provide pure rotational [1], pure translational [2], and a mixed rotational-translational motion of the platform with respect to the base [3]. An interesting 3-DOF PMs is the 3-UPU translational parallel manipulator that provides a pure translational motion of the platform with respect to the base: It was presented by Tsai in [4] and later used as a benchmark to investigate many important issues related to singularities, workspace analysis, and joint clearance influence on the platform accuracy. Hereafter this mechanism will be called 3-UPU TPM.

One of the most important characteristics of parallel manipulators is the stiffness, which

is strictly related with positioning accuracy and payload capability. Since the stiffness of the 3-UPU TPM (for a given configuration of the manipulator) depends on the moment applied by the platform to the legs, and this moment can be decomposed in two moments, namely a torque around the direction of the leg and a bending moment around a direction orthogonal to the direction of the leg itself, it is interesting to study the variation of these moments in the whole workspace in order to select the best location of a given trajectory that is characterized by the highest stiffness of the manipulator. A few authors have studied the stiffness of the 3-UPU TPM. In [5], the stiffness of four 3-DOF TPM architectures are compared, under the assumption that the manipulator compliance is due to the actuator compliance. For instance, the RUU architecture with rotational actuators is stiffer than the UPU architecture with linear actuators. However, if the manipulator compliance is due to link elasticity, Goldsmith [6] demonstrated that the 3-UPU TPM architecture with linear actuators is stiffer than that with revolute joints.

This paper will focus on 3-UPU TPM shown in Fig.1. The distribution of the torsion and the flexion on the legs of the manipulator in the workspace is investigated and a procedure for finding the best location of a given trajectory in the manipulator workspace is presented.

The paper is organized as follows: section 2 recalls, by referring to a previous paper of the authors [7], the kinetostatic analysis of the 3-UPU manipulator. Section 3 presents an analysis of the torsion and the flexion of the legs in the whole workspace, while section 4 reports a case study. Finally some conclusions are reported.

2. KINETOSTATIC ANALYSIS OF THE 3-UPU TPM

A schematic of a 3-UPU TPM is shown in Fig. 2. It features a moving 3-DOF translational platform connected to a fixed base by three extensible legs of type UPU, where U and P are for universal and actuated prismatic pair respectively. Each U joint comprises two revolute pairs with intersecting and perpendicular axes, centered at point B_i , $i=1,2,3$, in the base and at point A_i , $i=1,2,3$, in the platform. In order to have a pure translational motion of the platform, two conditions have to be fulfilled for each leg [4,8,9]:

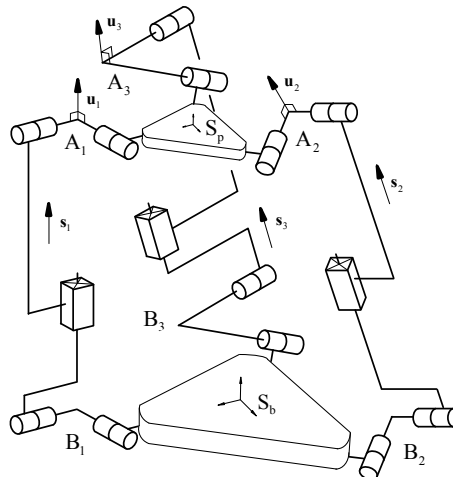


Figure 1. The 3-UPU Translational Parallel Manipulator.

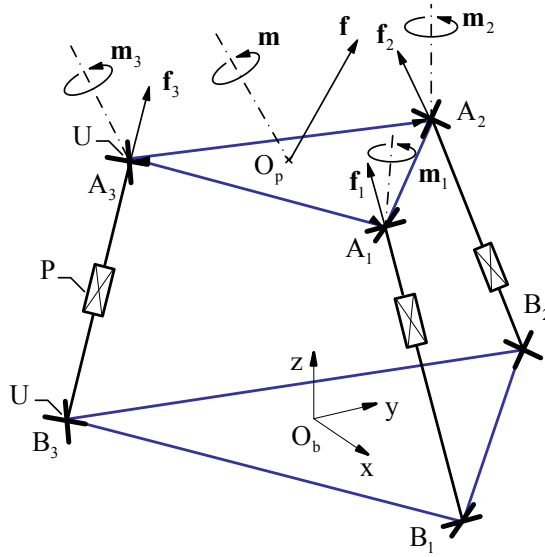


Figure 2. Different forces and moments applied on the platform.

- the axes of the two intermediate revolute pairs must be parallel to each other,
- the axes of the two ending revolute pairs must be parallel to each other.

The geometry studied in this paper of the manipulator is the following:

- the axes of the revolute pairs in the base/platform two- by-two intersect at three points,
- points A_i and B_i , $i=1,2,3$, form an equilateral triangle respectively.

With reference to Fig. 2, a system S_b (O_b, x, y, z) fixed to the base with origin O_b (the centre of the circle with radius b defined by the points B_i , $i=1,2,3$) is chosen. Axes x and y are on the plane defined by points B_i , $i=1,2,3$, with x axis through point B_1 . z pointing from the base to the platform and y axis is taken according to the right hand rule.

The balance of the platform involves the following wrenches:

- $\tau_{ext} = \{\mathbf{f}, \mathbf{m}\}$, the wrench of the external load applied on the platform at point O_p (the centre of the circle with radius p defined by the points A_i , $i=1,2,3$).
- $\tau_{ai} = \{\mathbf{f}_i, \mathbf{m}_i\}$, the wrenches provided by each of the i -th leg expressed at point A_i .

The wrench provided by the i -th leg to the platform is defined by the following equations [7]:

$$\begin{cases} \mathbf{f}_i = f_i \mathbf{s}_i \\ \mathbf{m}_i = m_i \mathbf{u}_i \end{cases} \quad i = 1, 2, 3 \quad (1)$$

where \mathbf{s}_i is a unit vector of the direction of the i -th leg and \mathbf{u}_i is a unit vector orthogonal to the cross of the universal joint as shown in Fig. 1.

The equilibrium of the platform expressed at point O_p is given in the following matrix form [7]:

$$\begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \mathbf{s}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{r}_{p1} \times \mathbf{s}_1 & \mathbf{r}_{p2} \times \mathbf{s}_2 & \mathbf{r}_{p3} \times \mathbf{s}_3 & \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{m} \end{bmatrix} \quad (2)$$

where:

$$\mathbf{r}_{pi} = \overrightarrow{O_p A_i} \quad i = 1, 2, 3 \quad (3)$$

When the manipulator is far from singularity, i.e., the determinant of the Jacobian matrix (6x6) defined by Eqn. (2) is different from zero, the module of the moment applied on the i-th leg, m_i can be computed as:

$$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} -\mathbf{U}^{-1} \mathbf{R} \mathbf{S}^{-1} & \mathbf{U}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{m} \end{bmatrix} \quad (4)$$

where \mathbf{S} , \mathbf{R} and \mathbf{U} are 3x3 matrices defined as follow :

$$\mathbf{S} = [\mathbf{s}_1 \quad \mathbf{s}_2 \quad \mathbf{s}_3] \quad (5)$$

$$\mathbf{R} = [\mathbf{r}_{p1} \times \mathbf{s}_1 \quad \mathbf{r}_{p2} \times \mathbf{s}_2 \quad \mathbf{r}_{p3} \times \mathbf{s}_3] \quad (6)$$

$$\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] \quad (7)$$

3. TORSION AND FLEXION OF THE LEGS OF THE 3-UPU TPM

In this section, the torsion and flexion of the legs of the manipulator due to the torque and the bending moment respectively are determined. The platform, the base and the universal joints are considered as rigid body.

The total moment m_i , $i=1,2,3$, applied by the platform to the i-th leg which is already computed by Eqn. (4) can be decomposed in two moments, m_{ti} , the torque about \mathbf{s}_i and m_{bi} , the bending moment about a direction \mathbf{b}_i orthogonal to the plane defined by the vectors \mathbf{s}_i and \mathbf{u}_i as shown in Fig. 3.

$$m_i \mathbf{u}_i = m_{ti} \mathbf{s}_i + m_{bi} \mathbf{b}_i \quad i = 1, 2, 3 \quad (8)$$

where the torque m_{ti} and the bending moment m_{bi} can be expressed as follow:

$$m_{ti} = m_i \cos \phi_i \quad i = 1, 2, 3 \quad (9)$$

$$m_{bi} = m_i \sin \phi_i \quad i = 1, 2, 3 \quad (10)$$

where ϕ_i is the angle defined by the unit vectors \mathbf{s}_i and \mathbf{u}_i as shown in Fig. 3.

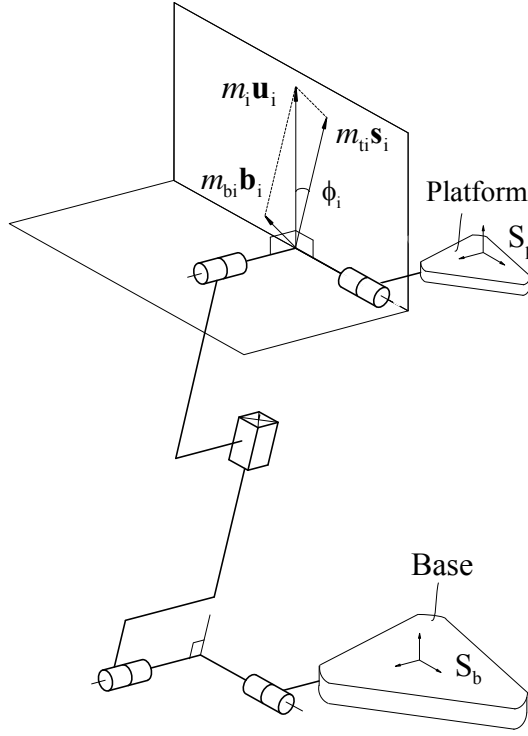


Figure 3. Moments applied on the i -th leg of the 3-UPU TPM.

According to Eqs. (9,10), the rate between the bending moment and the torque applied on the i -th leg, which does not depend to the module of the moment m_i , (it does not depend to the wrench of the external load applied on the platform), is given by:

$$\frac{m_{bi}}{m_i} = \frac{\sin \phi_i}{\cos \phi_i} \quad i = 1, 2, 3 \quad (11)$$

By computing the bending moment/torque rate, in the workspace, three special sections S_1 , S_2 and S_3 (they are defined geometrically by the help of Fig. 4) of the working space are found:

$$S_1: m_{b1} = 0; m_{b2}/m_{t2} = m_{b3}/m_{t3} \quad (12)$$

$$S_2: m_{b2} = 0; m_{b1}/m_{t1} = m_{b3}/m_{t3} \quad (13)$$

$$S_3: m_{b3} = 0; m_{b1}/m_{t1} = m_{b2}/m_{t2} \quad (14)$$

Figure 4 presents a view from the top of the manipulator (only the revolute pairs on the base and the platform are represented for clarity) and the three section defined above. The intersection of the sections S_1 , S_2 and S_3 with the plane (x,y) corresponds to three lines L_{S1} , L_{S2} and L_{S3} respectively.

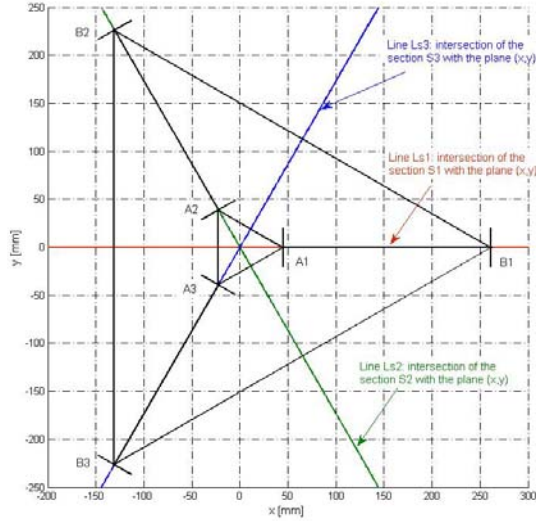


Figure 4. View from the top of the 3-UPU TPM and the three section S_1 , S_2 and S_3 .

These sections are used because of their properties (they are a plane of symmetry and the flexion is null in one of the three leg as given in Eqs. (12,13,14)), to find the location (best position of the platform) where the torque and the bending moments have the lowest values. In other words, where the leg torsion and leg flexion are the lowest.

Torsion of the legs

In this section, the analytic expression of the torsion of the legs of the 3-UPU TPM is presented. This expression is used in order to find the best position of the platform which corresponds to the lowest torsion of the legs.

According to Eqs. (4,9), the relation between the wrench of the external load applied on the reference point of the platform O_p and the torques applied on the legs is the following:

$$\begin{bmatrix} m_{t1} \\ m_{t2} \\ m_{t3} \end{bmatrix} = \mathbf{H}_1 \begin{bmatrix} \mathbf{f} \\ \mathbf{m} \end{bmatrix} \quad (15)$$

where \mathbf{H}_1 is a 3x6 matrix which depends on the position of the reference point O_p of the platform and the geometry of the manipulator.

The torsion θ_i of the i -th leg (not considering the elasticity of the joints) is given by:

$$\theta_i = \frac{l_i m_{ti}}{GI_0} \quad i = 1, 2, 3 \quad (16)$$

where I_0 , G and l_i are respectively the polar moment, the Coulomb module and the length of the i -th leg.

Equation (16), taking into account Eqn. (15), and extended to the three legs, becomes:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \mathbf{C}_1 \begin{bmatrix} \mathbf{f} \\ \mathbf{m} \end{bmatrix} \quad (17)$$

where:

$$\mathbf{C}_1 = [\mathbf{K}_1 | \mathbf{T}_1] \quad (18)$$

In Eqn. (18), \mathbf{K}_1 and \mathbf{T}_1 are a 3x3 matrices which represent respectively the influence of the external forces and the external moments on the torsion of the legs.

In order to compute the torsion of the legs when the point O_p moves on the section S_1 , S_2 and S_3 , the maximum or the minimum or the product of the singular values of the matrix \mathbf{C}_1 can be taken as criterion. However, the components of this matrix does not have the same units [10]. To solve the problem, some authors used a concept of the characteristic length which is related to the size of the platform [11]. Thus, Eqn. (17) is re-defined as follow:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \left[\frac{1}{L} \mathbf{K}_1 | \mathbf{T}_1 \right] \begin{bmatrix} \mathbf{f} \\ \mathbf{m} \end{bmatrix} \quad (19)$$

where L is the characteristic length that is taken equal the radius p of the circle that belongs the center of the universal joint connected to the platform A_i , $i=1,2,3$.

If the wrench of the external load corresponds only to a forces, Eqn. (19) becomes:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \frac{1}{L} \mathbf{K}_1 \mathbf{f} \quad (20)$$

In general, the determinant of (\mathbf{K}_1/L) is different from zero. In other words, where the three components of the external force \mathbf{f} applied to the platform, F_1 , F_2 and F_3 respectively along x , y and z axes of S_b belongs to a unit sphere, the torsions of the three legs θ_1 , θ_2 and θ_3 belong to an ellipsoid.

For the chosen geometry of the manipulator, we found that the determinant of (\mathbf{K}_1/L) is equal to zero in all the workspace. Thus, the ellipsoid becomes to an ellipse as shown in Fig. 5, i.e., there are two out of the three torsions that are dependent in all the workspace. This can be an advantage of this geometry of the manipulator. For example, in order to minimize three torsions, it is sufficient to minimize only two of them.

If the reference point O_p of the platform belongs the section S_1 defined above, and the components F_2 is equal zero, it can be obtained:

$$m_{11} = 0 \quad (21)$$

$$m_{12} = -m_{13} \quad (22)$$

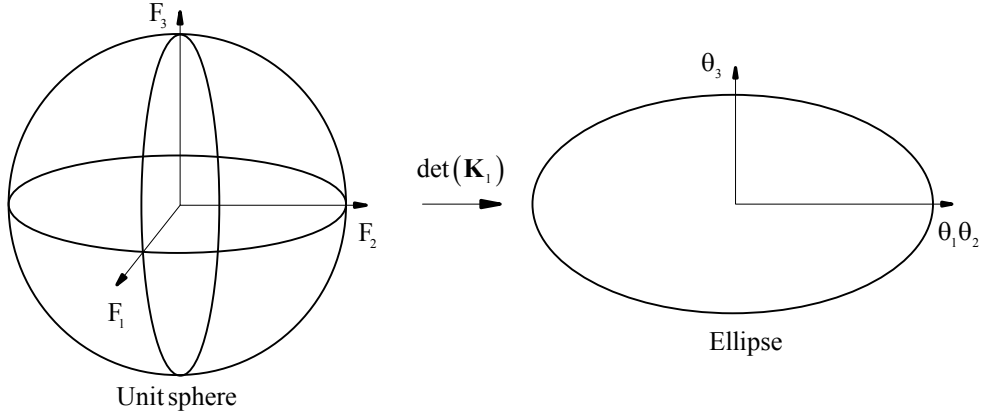


Figure 5. Shape of the distribution of the torsions θ_1 and θ_2 versus the torsion θ_3 .

According to Eqs. (21,22), the best position of the point O_p in the section S_1 is obtained where the torsion of the second leg is the lowest. By symmetry, the best position of the point O_p in the sections S_2 and S_3 is obtained where the torsion of the first and second leg is the lowest.

Flexion of the legs

In this section, the analytic expression of the flexion of the legs of the 3-UPU TPM is presented. This expression is used in order to find the best position of the platform which corresponds to the lowest flexion of the legs.

According to Eqs. (4,10), the relation between the wrench of the external load applied on the point O_p of the platform and the bending moments applied by the platform to the legs is given by:

$$\begin{bmatrix} m_{b1} \\ m_{b2} \\ m_{b3} \end{bmatrix} = \mathbf{H}_2 \begin{bmatrix} \mathbf{f} \\ \mathbf{m} \end{bmatrix} \quad (23)$$

where \mathbf{H}_2 is a 3x6 matrix which depends to the position of the point O_p and on the geometry of the manipulator.

The bending moment m_{bi} is constant for all the sections of the i -th leg (the shear stress is zero). Thus, the flexion ϕ_i of the i -th leg can be written as:

$$\phi_i = \frac{l_i m_{bi}}{EI} \quad i = 1, 2, 3 \quad (24)$$

where E and I are respectively the Young modulus and the moment of inertia of the section of the i -th leg.

Equation (24), taking into account Eqn. (23), extended to the three legs and with the introduction of the characteristic length L , becomes:

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix} = \mathbf{C}_2 \begin{bmatrix} \mathbf{f} \\ \mathbf{m} \end{bmatrix} \quad (25)$$

where:

$$\mathbf{C}_2 = \left[\frac{1}{L} \mathbf{K}_2 \middle| \mathbf{T}_2 \right] \quad (26)$$

In Eqn. (26), \mathbf{K}_2 and \mathbf{T}_2 are 3x3 matrices which represent respectively the influence of the external force and the external moment on the flexion of the legs.

In order to compute the flexion of the legs when the point O_p moves on the section S_1 , S_2 and S_3 , the maximum or the minimum or the product of the singular values of the matrix \mathbf{C}_2 can be taken as criterion. However, where the point O_p belongs to the sections S_1 , the bending moment applied to the first leg is equal to zero:

$$m_{b1} = 0 \quad (27)$$

Thus, the best position of the reference point O_p in the section S_1 is obtained where the flexion of the second and the third leg are the lowest.

As a general result, the best location of the reference point of the platform in the section S_1 is found by satisfying the following conditions:

- the torsion of the second leg is the lowest.
- the flexion of the second and the third leg are the lowest.

4. CASE STUDY

This section reports the distribution of the flexions and the torsion of the three legs of the 3-UPU translational parallel manipulator in the section S_1 of the workspace in order to find their lowest values.

The given data are:

- the radii of the two circles that belong the centers of the universal joint connected respectively to the base and to the platform respectively: $b = 261$ mm, $p = 45$ mm.
- The wrench of the external load corresponds only to an external forces.

For each leg (circular section):

- the external radius $R_{ext} = 8$ mm; the internal radius $R_{int} = 5.5$ mm;
- Young modulus E and Coulomb modulus of the Aluminium G are respectively $E = 69000$ N/mm²; $G = 26000$ N/mm².

The distribution of the torsion and the flexion of the legs of the manipulator in the section S_1 is shown in Fig. 6.a,b. The lines shown in Fig. 6.a and 6.b represent a trajectories in the section S_1 which have the same value of the torsion and the flexion respectively.

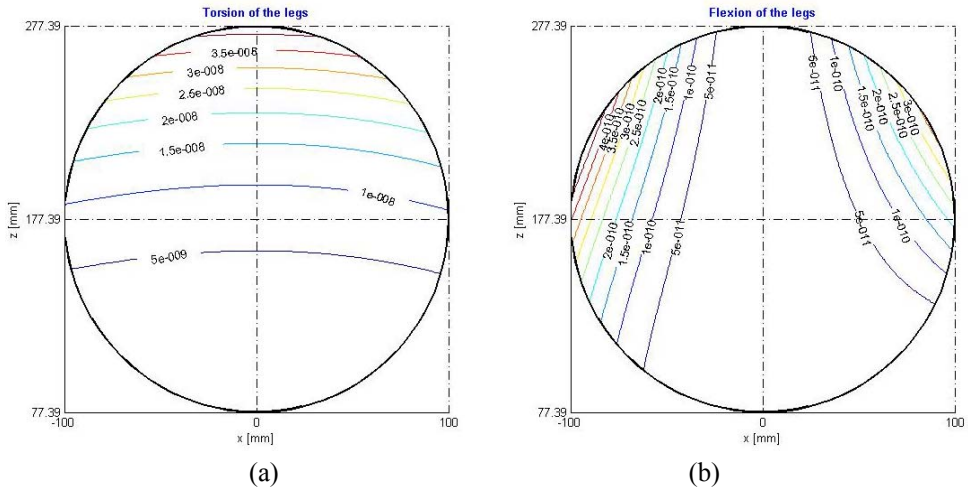


Figure 6. Distribution of the torsion and the flexion of the legs in the section S_1 .

According to these figures, it is easy to select the best position of the platform in the section S_1 which correspond to the lowest torsion and flexion of the second leg and the lowest flexion of the third legs.

5. CONCLUSION

An investigation has been presented to show the distribution of the torsion and the flexion on the legs of the 3-UPU translational manipulator in a section of its workspace. The best location of a given trajectory in the manipulator workspace, that is where torsion and flexion of the legs is the lowest, corresponds to the lowest value of the z coordinate of the reference point of the platform. The presented procedure can be used, for instance, for the design of the manipulator with given stiffness properties and the selection of the best positioning of a given trajectory in the manipulator workspace.

Acknowledgements

The financial support of MIUR, INAIL, and AERTECH-Lab is gratefully acknowledged.

REFERENCES

- [1] Gosselin, C., Angeles, J., 1989. "The optimum kinematic design of a spherical three degree-of-freedom parallel manipulator". *ASME Journal of Mechanisms, Transmission and Automation in Design*, vol. 111, pp. 202-207.
- [2] Hervé, J.M., Sparacino, F., 1991. "Structural synthesis of parallel robots generating spatial translation". In: *Fifth ICAR International Conference on Advanced Robotics*, pp. 808-813, Pisa, Italy.
- [3] Po-hua, Y., Waldron, K., Orin, D.E., 1996. "Kinematic of a three-degree-of-freedom motion platform for a low-cost driving simulator". *Advances in Robot Kinematics*, pp. 89-98, Kluwer Academic Publishers.

- [4] Tsai, L.W., 1996. "Kinematics of three-degrees of freedom platform with three extensible limbs". *Recent advances in robot kinematics*, pp.401-410, Kluwer, Dordrecht.
- [5] Tsai, L.W., Joski, S., 2001. "Comparison study of architectures of four 3 degree-of-freedom translational parallel manipulators" In *Proc. IEEE Int. Conf. Robotics Automation*, Seoul, Korea, 2001.
- [6] Goldsmith, P.B., 2002. "Kinematics and Stiffness of a Translational Parallel Symmetrical 3-UPU Manipulator" *Proceedings of the 2002 IEEE International Conference on Robotics & Automation*, Washington, DC May 2002
- [7] Wolf, A., Shoham, M., 2006. "Screw theory tools for the synthesis of the geometry of a parallel robot for a given instantaneous task". *Mechanism and Machine Theory*, vol. 41, pp. 656-670.
- [8] Di Gregorio, R., Parenti-Castelli, V., 1998. "A Translational 3-DOF Parallel Manipulator". *Advances in Robot Kinematics: Analysis and Control*, pp. 49-58.
- [9] Di Gregorio, R., Parenti-Castelli, V., 2002. "Mobility analysis of the 3-UPU parallel mechanism assembled for a pure translational motion". *ASME Transactions, Journal of Mechanical Design*, vol. 124, pp. 259-264.
- [10] Duffy, J., 1990. "The Fallacy of Modern Hybrid Control Theory that is Based on Orthogonal Complements' of twist and Wrench Spaces". *Journal of Robotic Systems*, vol. 7, pp. 139-144.
- [11] Angeles, J., 2002. *Fundamentals of Robotic Mechanical Systems. Theory, Methods, and Algorithms*, 2nd ed., Springer-Verlag, New York.

SYNTHESIS OF THE SPHERICAL FOUR-BAR LINKAGE FUNCTION GENERATOR FOR SIX PRECISION POINTS

Carlo Innocenti

Department of Mechanical and Civil Engineering,

University of Modena and Reggio Emilia, Italy

E-mail: carlo.innocenti@unimore.it

Abstract. *A procedure is presented to solve the dimensional synthesis of the spherical four-bar linkage function generator, once six precision points are specified. All feasible linkages result from solving a univariate seventh-order algebraic equation. Differently from existing techniques, the proposed approach does not introduce extraneous solutions, i.e., any of the resulting spherical four-bar linkages can be assembled at all six prescribed precision points. A numerical example shows application of the presented procedure to a case study.*

Keywords: *dimensional synthesis, spherical mechanism, function generator*

1. INTRODUCTION

The dimensional synthesis of function generators is a classical topic in kinematics [1]. A function generator is a one-degree-of-freedom linkage that exhibits a prescribed relationship (the input-output relationship) between the absolute positions of two links directly connected to the frame (the input and output links). Different types of requirements can be imposed on a function generator: if the input-output relationships involves an infinite set of positions for the input and output links, then there generally exist only approximate solutions, which result from optimization processes. Conversely, when a linkage has to comply with a finite set of position pairs (the so-called precision points) for the input and output links, then an exact solution might exist depending on the type of linkage at hand, the number of precision points, and their actual values.

Obviously, complex contraptions are potentially less appealing as function generators than structurally simple linkages such as, for instance, the planar and spherical four-bar linkages. The latter, in particular, is suitable to connect two concurrent-axis shafts, which are themselves the input and output links of the function generator.

The dimensional synthesis of the spherical four-bar linkage function generator (S4BLFG) has been dealt with in the technical literature for a number of different tasks. Exact solutions have been proposed for different numbers of precision points: three [1-2], four [2-4], five [4-7], and six [4, 8-10]. Optimized solutions have been presented in [11-13].

It is worth observing that, when the angle between the connected shafts is a free design parameter too, a S4BLFG is defined by six parameters, as many as the links of the linkage, plus two offset angles. Therefore, six is the maximum number of arbitrarily chosen precision points that can be specified for a S4BLFG.

With regard to the dimensional synthesis of the S4BLFG for six precision points, an algorithm has been recently proposed in [10] that is a step forward with respect to the previously known techniques. According to [10], the dimensional synthesis of a S4BLFG

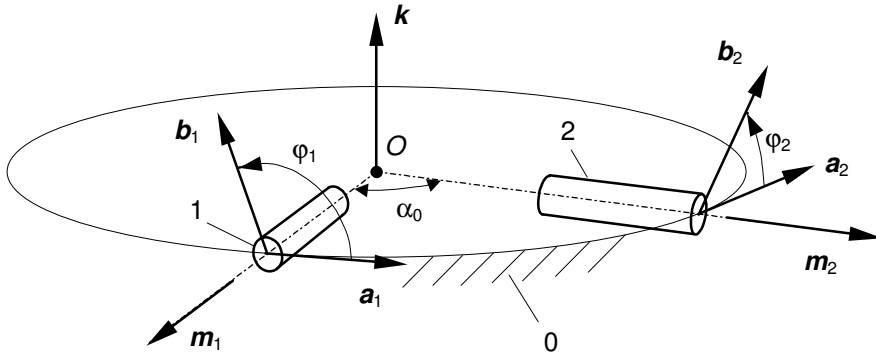


Figure 1. The input and output links of the S4BLFG.

can be carried out by solving a tenth-order polynomial equation and, for any of its real roots, second-order polynomial equations. Therefore, a maximum number of twenty candidate solutions can be determined. Unfortunately, not all of them are acceptable, because the solving procedure described in [10] introduces extraneous solutions, i.e., solutions that do not satisfy the prescribed precision points. All candidate solutions must then be scrutinized *a posteriori* in order to discard the extraneous ones (there are thirteen of them).

This paper presents a new procedure to solve the dimensional synthesis of the spherical four-bar linkage function generator for six precision points. The proposed procedure does not introduce extraneous solutions. It leads to a seventh-order algebraic equation that is pivotal in determining all possible answers to the considered dimensional synthesis problem. For any real root of this algebraic equation, if all dimensional parameters of the corresponding spherical four-bar linkage turn out to be real, then a real linkage exists that can be assembled at all prescribed precision points.

Although the proposed algorithm yields prospective linkages that comply with the specified precision points, this does not necessarily imply that any of such linkages can satisfactorily perform as a function generator. Further input data and inspection are needed in order to rule out chances of circuit, branch, and order defects [14]. Nevertheless, these issues are beyond the scope of this paper.

A numeric examples shows application of the proposed procedure to a case study.

2. NOMENCLATURE

With reference to a spherical four-bar linkage function generator (S4BLFG), the axes of the revolute pairs that connect the input and output links to the frame meet at point O , the center of the spherical linkage (see Fig. (1)). The distinction between input and output links is not relevant for the present study. Therefore, these two links will be arbitrarily identified by labels 1 and 2, irrespective of their roles. Furthermore, label 0 and 3 will denote, in the order, the frame and coupler of the S4BLFG. The direction of the axis of the revolute pair between link i ($i=1,2$) and the frame is defined by unit vector m_i , whose orientation is

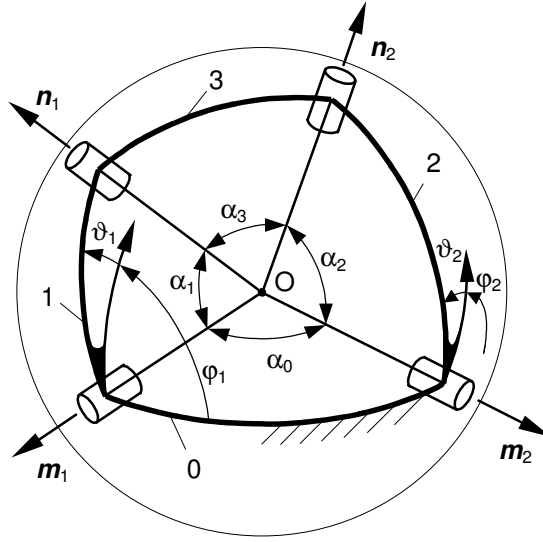


Figure 2. The S4BLFG.

consistent with the arbitrarily chosen direction of the positive rotations of link i relative to the frame (see Fig. (1)).

By assuming that \mathbf{m}_1 and \mathbf{m}_2 are not parallel, a unit vector \mathbf{k} is introduced as defined by

$$\mathbf{k} = \frac{\mathbf{m}_1 \times \mathbf{m}_2}{\|\mathbf{m}_1 \times \mathbf{m}_2\|} \quad (1)$$

The angle α_0 between unit vectors \mathbf{m}_1 and \mathbf{m}_2 – an unknown design parameter of the S4BLFG – is indirectly provided by

$$\cos \alpha_0 = \mathbf{m}_1 \cdot \mathbf{m}_2; \quad \sin \alpha_0 = \mathbf{m}_1 \times \mathbf{m}_2 \cdot \mathbf{k} \quad (2)$$

Angle α_0 can take any value in the range $]0, \pi[$.

Unit vectors \mathbf{a}_i ($i=1,2$) fixed to the frame are now considered as defined by

$$\mathbf{a}_i = \mathbf{k} \times \mathbf{m}_i \quad (i=1,2) \quad (3)$$

A unit vector \mathbf{b}_i orthogonal to unit vector \mathbf{m}_i is then attached to link i ($i=1,2$; see Fig. (1)). The angular position of link i relative to the frame is provided by the angle φ_i between \mathbf{a}_i and \mathbf{b}_i . More precisely, angle φ_i ($-\pi < \varphi_i \leq \pi$) is indirectly defined by

$$\cos \varphi_i = \mathbf{a}_i \cdot \mathbf{b}_i; \quad \sin \varphi_i = \mathbf{a}_i \times \mathbf{b}_i \cdot \mathbf{m}_i \quad (i=1,2) \quad (4)$$

Figure (2) shows a schematic of the S4BLFG. A unit vector \mathbf{n}_i is introduced parallel to the axis of the revolute pair between link i ($i=1,2$) and the coupler 3. The orientation of \mathbf{n}_i is

such as to make the angle α_i between vectors \mathbf{m}_i and \mathbf{n}_i fall into interval $[0, \pi/2]$; therefore its sine is non-negative and its cosine – non-negative too – is provided by

$$\cos \alpha_i = \mathbf{m}_i \cdot \mathbf{n}_i \quad (i = 1, 2) \quad (5)$$

The plane that contains the axes of the revolute pairs that connect link i ($i=1,2$) to the frame and the coupler is not necessarily parallel to unit vector \mathbf{b}_i . The offset angle ϑ_i is therefore introduced to account for this possibility. Angles ϑ_i ($i=1,2$) are unknown design parameters of the S4BLFG. They can take any value in the interval $[-\pi, \pi]$. If $\varphi_i + \vartheta_i$ were equal to zero, then unit vector \mathbf{n}_i would be parallel to the plane defined by unit vectors \mathbf{m}_1 and \mathbf{m}_2 , with a non-negative component along $\mathbf{k} \times \mathbf{m}_i$.

The only kinematically-relevant dimension of the coupler of the S4BLFG is angle α_3 (see Fig. (2)). It is the last unknown parameter of the S4BLFG. Angle α_3 can always be considered as non-negative, with a magnitude smaller than π radians. It is indirectly defined by

$$\cos \alpha_3 = \mathbf{n}_1 \cdot \mathbf{n}_2 \quad (6)$$

Collectively there are six design parameters for the considered dimensional synthesis problem regarding the S4BLFG: they are the link angular dimensions α_i ($i=0,...,3$), together with offset angles ϑ_i ($i=1,2$).

As already mentioned, the value of angle α_0 is conventionally bound to interval $]0, \pi[$, angles α_1 and α_2 are restricted to interval $[0, \pi/2]$, whereas angle α_3 is confined to interval $[0, \pi[$. No such limitations exist for angles ϑ_i ($i=1,2$), which can range throughout interval $[-\pi, \pi]$. These restrictions do not prevent any possible solution of the dimensional synthesis problem to be found. Actually, they conveniently rule out the chance of counting the same solution more than once since, without them, the same S4BLFG could be described by more than one set of angles α_i ($i=0,...,3$) and ϑ_i ($i=1,2$).

3. THE COMPATIBILITY EQUATIONS

In order to determine the six design parameters of the S4BLFG, six equations are written, one for each of the prescribed precision points. Figure (3) shows the S4BLFG at the configuration that corresponds to the j -th precision point ($j=1,...,6$), which in turn is provided by the values φ_{1j} and φ_{2j} for angles φ_1 and φ_2 respectively.

A Cartesian reference frame $Oxyz$ is introduced with origin at the center O of the spherical linkage, x -axis with the same direction and orientation as unit vector \mathbf{m}_1 , z -axis with the same direction and orientation as unit vector \mathbf{k} .

At the j -th precision point ($j=1,...,6$), the components of unit vector \mathbf{n}_1 in reference frame $Oxyz$ are given by

$$\mathbf{n}_{1j} = \mathbf{R}_x(\vartheta_1 + \varphi_{1j}) \begin{pmatrix} u_1 \\ v_1 \\ 0 \end{pmatrix} \quad (7)$$

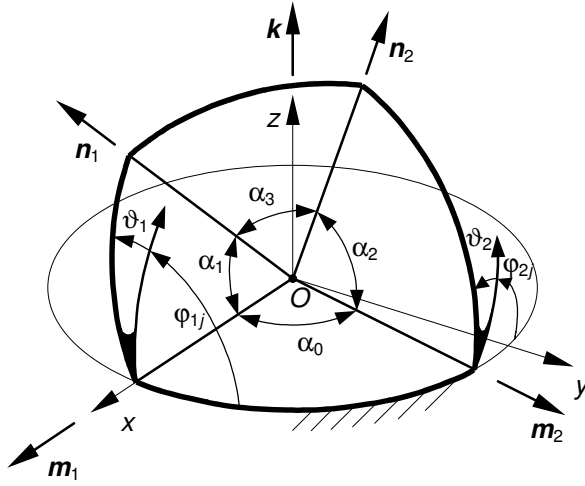


Figure 3. The S4BLFG at a generic precision point.

In this equation \mathbf{R}_x is a 3×3 orthogonal matrix that depends on angle $\vartheta_1 + \varphi_{1j}$. For a generic angle ξ , matrix $\mathbf{R}_x(\xi)$ is given by

$$\mathbf{R}_x(\xi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \xi & -\sin \xi \\ 0 & \sin \xi & \cos \xi \end{pmatrix} \quad (8)$$

The components of the vector on the right-hand side of Eqn. (7) stem from the ensuing shorthand notation for the cosine and sine of angles α_i ($i=0, \dots, 3$)

$$u_i = \cos \alpha_i; \quad v_i = \sin \alpha_i \quad (i = 0, \dots, 3) \quad (9)$$

Similarly, at the j -th precision point the components of unit vector \mathbf{n}_2 in reference frame $Oxyz$ are given by

$$\mathbf{n}_{2j} = \mathbf{R}_z(\alpha_0) \mathbf{R}_x(\vartheta_2 + \varphi_{2j}) \begin{pmatrix} u_2 \\ v_2 \\ 0 \end{pmatrix} \quad (10)$$

Matrix $\mathbf{R}_z(\alpha_0)$ on the right-hand side of this equation is the 3×3 orthogonal matrix defined by

$$\mathbf{R}_z(\alpha_0) = \begin{pmatrix} \cos \alpha_0 & -\sin \alpha_0 & 0 \\ \sin \alpha_0 & \cos \alpha_0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (11)$$

The six compatibility equations for the dimensional synthesis problem here considered are concisely given by

$$\mathbf{n}_{1j} \cdot \mathbf{n}_{2j} = u_3 \quad (j = 1, \dots, 6) \quad (12)$$

By adopting the additional shorthand notations

$$c_i = \cos \vartheta_i; \quad s_i = \sin \vartheta_i \quad (i = 1, 2) \quad (13)$$

$$E_{ij} = \cos \phi_{ij}; \quad F_{ij} = \sin \phi_{ij} \quad (i = 1, 2; \quad j = 1, \dots, 6) \quad (14)$$

Eqn. (12) is re-written in extended form as follows

$$\begin{aligned} & F_{1j}F_{2j}c_1c_2v_1v_2 + E_{1j}F_{2j}s_1c_2v_1v_2 + F_{1j}E_{2j}c_1s_2v_1v_2 + E_{1j}E_{2j}s_1s_2v_1v_2 \\ & + E_{1j}E_{2j}c_1c_2u_0v_1v_2 - F_{1j}E_{2j}s_1c_2u_0v_1v_2 - E_{1j}F_{2j}c_1s_2u_0v_1v_2 + F_{1j}F_{2j}s_1s_2u_0v_1v_2 \\ & - E_{2j}c_2u_1v_0v_2 + F_{2j}s_2u_1v_0v_2 + E_{1j}c_1u_2v_0v_1 - F_{1j}s_1u_2v_0v_1 + u_0u_1u_2 = u_3 \end{aligned} \quad (j = 1, \dots, 6) \quad (15)$$

The next section will show a procedure to solve the six equations (15) for the six unknowns α_i ($i=0, \dots, 3$) and ϑ_i ($i=1, 2$), which in turn appear in Eqn. (15) through their elementary trigonometric functions (see Eqns. (9) and (13)).

4. EQUATION SOLUTION

Elimination of unknown α_3 from the six-equation set (15) is straightforwardly carried out by subtracting the sixth equation from the previous five. The resulting five equations are

$$\sum_{\substack{i=1, \dots, 3 \\ j=1, \dots, 4}} G_{ijk} g_{ij} = 0 \quad (k = 1, \dots, 5) \quad (16)$$

Constant coefficients G_{ijk} ($i=1, \dots, 3$; $j=1, \dots, 4$; $k=1, \dots, 5$) in Eqn. (16) depend on the problem data only; their expressions – here omitted for the sake of brevity – simply result from the coefficients on the left-hand side of Eqn. (15). Quantities g_{ij} ($i=1, \dots, 3$; $j=1, \dots, 4$) are functional terms collectively defined by the ensuing matrix relation

$$\begin{pmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \end{pmatrix} = \begin{pmatrix} c_1c_2v_1v_2 & s_1c_2v_1v_2 & c_1s_2v_1v_2 & s_1s_2v_1v_2 \\ c_1c_2u_0v_1v_2 & s_1c_2u_0v_1v_2 & c_1s_2u_0v_1v_2 & s_1s_2u_0v_1v_2 \\ c_2(u_1v_0)v_2 & s_2(u_1v_0)v_2 & c_1(u_2v_0)v_1 & s_1(u_2v_0)v_1 \end{pmatrix} \quad (17)$$

The five unknowns in Eqn. (16) are α_i ($i=0, \dots, 2$) and ϑ_i ($i=1, 2$).

Replacement of Some Unknowns

On the bottom row of the 3×4 matrix on the right-hand side of Eqn. (17) some terms have been enclosed within parentheses. Now these terms are replaced by the ensuing expressions

$$u_1 v_0 = w_1 c_1 v_1; \quad u_2 v_0 = w_2 c_2 v_2 \quad (18)$$

Consequently, Eqn. (16) is re-written as

$$\sum_{\substack{i=1,...,3 \\ j=1,...,4}} G_{ijk} h_{ij} = 0 \quad (k=1,...,5) \quad (19)$$

with

$$\begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \end{pmatrix} = \begin{pmatrix} c_1 c_2 v_1 v_2 & s_1 c_2 v_1 v_2 & c_1 s_2 v_1 v_2 & s_1 s_2 v_1 v_2 \\ c_1 c_2 u_0 v_1 v_2 & s_1 c_2 u_0 v_1 v_2 & c_1 s_2 u_0 v_1 v_2 & s_1 s_2 u_0 v_1 v_2 \\ w_1 c_1 c_2 v_1 v_2 & w_1 c_1 s_2 v_1 v_2 & w_2 c_1 c_2 v_2 v_1 & w_2 s_1 c_2 v_1 v_2 \end{pmatrix} \quad (20)$$

Since angles α_1 and α_2 are supposed to belong to interval $[0, \pi/2]$, the product $v_1 v_2$ cannot vanish (see Eqn. (9)). Therefore Eqn. (19) can be divided by $v_1 v_2$, thus obtaining

$$\sum_{\substack{i=1,...,3 \\ j=1,...,4}} G_{ijk} p_{ij} = 0 \quad (k=1,...,5) \quad (21)$$

with

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} = \begin{pmatrix} c_1 c_2 & s_1 c_2 & c_1 s_2 & s_1 s_2 \\ c_1 c_2 u_0 & s_1 c_2 u_0 & c_1 s_2 u_0 & s_1 s_2 u_0 \\ w_1 c_1 c_2 & w_1 c_1 s_2 & w_2 c_1 c_2 & w_2 s_1 c_2 \end{pmatrix} \quad (22)$$

The five equations in equation set (21) still contain five unknowns, i.e., α_0 , w_i ($i=1,2$), and ϑ_i ($i=1,2$).

Elimination of unknowns w_1 and w_2

Four of the five equations in equation set (21) are now linearly solved for the functional terms p_{3k} ($k=1,...,4$), which are the only functional terms that contain unknowns w_1 and w_2 . This step results into

$$p_{3k} = \sum_{\substack{i=1,2 \\ j=1,...,4}} H_{ijk} p_{ij} \quad (k=1,...,4) \quad (23)$$

The values of constant coefficients H_{ijk} ($i=1,2; j=1,...,4; k=1,...,4$) in Eqn. (23) depend exclusively on constant coefficients G_{ijk} ($i=1,...,3; j=1,...,4; k=1,...,5$) of Eqn. (21), in turn defined by the problem data only.

The expressions in Eqn. (23) for p_{3k} ($k=1,...,4$) are now inserted into the equation of set (21) that has not been exploited to obtain Eqn. (23). The ensuing equation is thus obtained

$$\sum_{i,j,k=1,2} L_{ijk} U_i X_j Y_k = 0 \quad (24)$$

In this equation, quantities L_{ijk} ($i,j,k=1,2$) are constant coefficients that depend exclusively on coefficients G_{ijk} . Moreover, quantities U_i , X_j , and Y_j ($i,j,k=1,2$) are the components of the ensuing vectors

$$\mathbf{U} = \begin{pmatrix} 1 \\ u_0 \end{pmatrix}; \quad \mathbf{X} = \begin{pmatrix} c_1 \\ s_1 \end{pmatrix}; \quad \mathbf{Y} = \begin{pmatrix} c_2 \\ s_2 \end{pmatrix} \quad (25)$$

Equation (24) is a first condition that does not contain unknowns w_1 and w_2 . Other two of such conditions can be obtained by the exploiting the definition of quantities p_{3k} ($k=1,...,4$) (see Eqn. (22))

$$\begin{cases} p_{31}s_2 - p_{32}c_2 = 0 \\ p_{33}s_1 - p_{34}c_1 = 0 \end{cases} \quad (26)$$

If expressions in Eqn. (23) for p_{3k} ($k=1,...,4$) are inserted into Eqn. (26), the ensuing conditions are obtained

$$\sum_{i,j,k,r=1,2} M_{ijk} U_i X_j Y_k Y_r = 0 \quad (27)$$

$$\sum_{i,j,k,r=1,2} N_{ijk} U_i X_j X_k Y_r = 0 \quad (28)$$

Coefficients M_{ijk} and N_{ijk} ($i,j,k,r=1,2$) in Eqns. (27)-(28) are constant quantities that depend on the problem data only.

Equations (24), (27), and (28) together form a set of three equations that contain α_0 , ϑ_1 , and ϑ_2 as only unknowns.

Elimination of unknowns α_0 and ϑ_2

The three equations in three unknowns derived in the previous subsection are now further manipulated in order to obtain a univariate equation in the unknown ϑ_1 .

By assuming $c_2 \neq 0$ (i.e., $\vartheta_2 \neq \pi/2$), Eqns. (24), (27), and (28) are divided, respectively, by c_2 , c_2^2 , and c_2 , thus obtaining

$$\sum_{i,j,k=1,2} L_{ijk} U_i X_j T_k = 0 \quad (29)$$

$$\sum_{i,j,k,r=1,2} M_{ijkr} U_i X_j T_k T_r = 0 \quad (30)$$

$$\sum_{i,j,k,r=1,2} N_{ijkr} U_i X_j X_k T_r = 0 \quad (31)$$

Quantities T_k ($k=1,2$) in Eqn. (29)-(31) are the components of the ensuing vector

$$\mathbf{T} = \begin{pmatrix} 1 \\ t_2 \end{pmatrix} \quad (32)$$

with

$$t_2 = \tan \vartheta_2 \quad (33)$$

Equations (29)-(31) are now regarded as three equations in the unknowns u_0 and t_2 , with coefficients that depend on ϑ_1 . They are respectively re-written as follows

$$\sum_{\substack{i=0,1 \\ j=0,1}} P_{ij} u_0^i t_2^j = 0 \quad (34)$$

$$\sum_{\substack{i=0,1 \\ j=0,2}} Q_{ij} u_0^i t_2^j = 0 \quad (35)$$

$$\sum_{\substack{i=0,1 \\ j=0,1}} R_{ij} u_0^i t_2^j = 0 \quad (36)$$

In these equations, quantities P_{ij} ($i,j=0,1$) and Q_{ij} ($i=0,1; j=0,2$) depend linearly on c_1 and s_1 (see Eqn. (13)), whereas R_{ij} ($i,j=0,1$) depends quadratically on c_1 and s_1 . More specifically, P_{ij} and Q_{ij} are sums of terms that depend linearly either on c_1 or s_1 , whereas R_{ij} is a sum of terms that depend linearly on only one of the ensuing functional quantities: c_1^2 , $c_1 s_1$, s_1^2 .

At this point a set of eight auxiliary equations is formed. The first three of such equations are obtained by multiplying Eqn. (34) by 1, t_2 , and t_2^2 . The subsequent two auxiliary equations derive from multiplying Eqn. (35) by 1 and t_2 . Finally, the last three auxiliary equations result from multiplication of Eqn. (36) by 1, t_2 , and t_2^2 . The set of eight auxiliary equations is written in matrix form as follows

$$\mathbf{K} \mathbf{f} = \mathbf{0} \quad (37)$$

Quantity \mathbf{K} in Eqn. (37) is the ensuing 8×8 matrix

$$\mathbf{K} = \begin{pmatrix} P_{00} & P_{01} & 0 & 0 & P_{10} & P_{11} & 0 & 0 \\ 0 & P_{00} & P_{01} & 0 & 0 & P_{10} & P_{11} & 0 \\ 0 & 0 & P_{00} & P_{01} & 0 & 0 & P_{10} & P_{11} \\ Q_{00} & Q_{01} & Q_{02} & 0 & Q_{10} & Q_{11} & Q_{12} & 0 \\ 0 & Q_{00} & Q_{01} & Q_{02} & 0 & Q_{10} & Q_{11} & Q_{12} \\ R_{00} & R_{01} & 0 & 0 & R_{10} & R_{11} & 0 & 0 \\ 0 & R_{00} & R_{01} & 0 & 0 & R_{10} & R_{11} & 0 \\ 0 & 0 & R_{00} & R_{01} & 0 & 0 & R_{10} & R_{11} \end{pmatrix} \quad (38)$$

whereas \mathbf{f} is the eight-component vector defined by

$$\mathbf{f} = (1 \quad t_2 \quad t_2^2 \quad t_2^3 \quad u_0 \quad u_0 t_2 \quad u_0 t_2^2 \quad u_0 t_2^3)^T \quad (39)$$

Since the first component of vector \mathbf{f} is different from zero, Eqn. (37) cannot be satisfied by a vanishing vector \mathbf{f} , which in turn means that matrix \mathbf{K} has to be singular, i.e.

$$\det \mathbf{K} = 0 \quad (40)$$

This equation is the result of eliminating unknowns α_0 and ϑ_2 from the three-equation set formed by Eqns. (24), (27), and (28).

The Univariate Polynomial Equation

The dependence on c_1 and s_1 is linear for the first five rows of matrix \mathbf{K} , quadratic for the remaining three rows. This assertion simply ensues from the already-established dependence of quantities P_{ij} , Q_{ij} , and R_{ij} on c_1 and s_1 . Therefore the left-hand side of Eqn. (40) is a linear combination – through known, constant coefficients – of terms the like of $s_1^i c_1^j$, with $i \geq 0$, $j \geq 0$, $i+j=11$. If in Eqn. (40) any power of c_1 greater than one is replaced by one of the ensuing expressions

$$c_1^{2i} \rightarrow (1 - s_1^2)^i; \quad c_1^{2i+1} \rightarrow c_1 (1 - s_1^2)^i \quad (i = 1, \dots, 5) \quad (41)$$

then Eqn. (40) takes the ensuing form

$$A(s_1^2) s_1 + B(s_1^2) c_1 = 0 \quad (42)$$

In this equation, $A(s_1^2)$ and $B(s_1^2)$ are polynomials in s_1^2 whose coefficients depend exclusively on the problem data. Based on the mentioned dependence of Eqn. (40) on c_1 and s_1 , the maximum power of s_1^2 in the expressions of $A(s_1^2)$ and $B(s_1^2)$ should not exceed five. Actual computation shows that of $A(s_1^2)$ and $B(s_1^2)$ are third-order polynomials in s_1^2 .

Equation (42) is now multiplied by c_1

$$s_1 A(s_1^2) c_1 + B(s_1^2) c_1^2 = 0 \quad (43)$$

Then, in Eqn. (43), c_1^2 is replaced with the expression provided by the first of Eqn. (41) for $i=1$

$$(1 - s_1^2) B(s_1^2) + s_1 A(s_1^2) c_1 = 0 \quad (44)$$

Equations (44) and (42) are now re-written together as follows

$$\begin{pmatrix} (1 - s_1^2) B(s_1^2) & s_1 A(s_1^2) \\ s_1 A(s_1^2) & B(s_1^2) \end{pmatrix} \begin{pmatrix} 1 \\ c_1 \end{pmatrix} = \mathbf{0} \quad (45)$$

Equation (45) is satisfied only if the 2×2 matrix on its left-hand side is singular, i.e.

$$(1 - s_1^2) [B(s_1^2)]^2 - s_1^2 [A(s_1^2)]^2 = 0 \quad (46)$$

Based on the order of polynomials $A(s_1^2)$ and $B(s_1^2)$, Eqn. (46) is of degree seven in s_1^2 . It can be re-written in the ensuing form

$$\sum_{i=0}^7 C_i s_1^{2i} = 0 \quad (47)$$

Coefficients C_i ($i=0, \dots, 7$) in Eqn. (47) are constant quantities that depend exclusively on the data of the dimensional synthesis problem. In the complex domain, Eqn. (47) is satisfied by seven values of s_1^2 , which can be numerically computed with the desired approximation.

Back Substitution

The spherical four-bar linkage function generators that comply with the prescribed six precision points can only originate from the real roots of Eqn. (48) that fall into interval $[0, 1]$. For the generic of these roots the value s_1 is yielded by

$$s_1 = \sqrt{s_1^2} \quad (49)$$

The corresponding value of c_1 stems from Eqn. (42)

$$c_1 = -s_1 \frac{A(s_1^2)}{B(s_1^2)} \quad (50)$$

Based on s_1 and c_1 , the value of offset angle ϑ_1 – one of the six design parameters of the S4BLFG – can be determined in the range $[0, \pi]$.

The already-found values of s_1 and c_1 make matrix \mathbf{K} defined by Eqn. (38) become singular. Seven out of the eight scalar equations in Eqn. (37) can be therefore linearly

Table 1. Ranges of the design parameters (rad).

Parameter	α_0	α_1	α_2	α_3	ϑ_1	ϑ_2
From back substitution	$]0, \pi[$	$\left]-\frac{\pi}{2}, \frac{\pi}{2}\right]$	$\left]-\frac{\pi}{2}, \frac{\pi}{2}\right]$	$[0, \pi[$	$[0, \pi[$	$\left]-\frac{\pi}{2}, \frac{\pi}{2}\right]$
After normalization	$]0, \pi[$	$\left[0, \frac{\pi}{2}\right]$	$\left[0, \frac{\pi}{2}\right]$	$[0, \pi[$	$]-\pi, \pi]$	$]-\pi, \pi]$

solved for the last seven components of vector \mathbf{f} (see Eqn. (39)). As a by-product of this process, the numeric values of u_0 and t_2 become available too.

If u_0 falls outside interval $[-1, 1]$, no real S4BLFG can be found for the considered root of Eqn. (47). Otherwise the value of design parameter α_0 can be determined in the range $]0, \pi[$ (see the first relation in Eqn. (9), for $i=0$).

Thanks to Eqn. (33), the value of t_2 allows determination of design parameter ϑ_2 in the range $]-\pi/2, \pi/2]$.

In order to compute design parameters α_1 and α_2 , the definitions in Eqn. (22) of p_{31} and p_{33} are preliminarily considered together with Eqn. (18). Elimination of quantities w_i ($i=1, 2$) from Eqns. (18) and (22) yields the ensuing relations

$$\tan \alpha_1 = \frac{v_0 c_2}{p_{31}}; \quad \tan \alpha_2 = \frac{v_0 c_1}{p_{33}} \quad (51)$$

Once quantities p_{31} and p_{33} have been computed by Eqn. (23), Eqn. (51) allows determination of α_1 and α_2 in the range $]-\pi/2, \pi/2]$.

Finally, any of the six conditions in Eqn. (15) yields the value of u_3 . If the absolute value of u_3 is not greater than one, design parameter α_3 can be determined in the range $[0, \pi[$ (see the first relation in Eqn. (9), for $i=3$).

Normalization of the Design Parameters

Whenever the procedure outlined in the previous subsection can be adhered to up to its ending point, a S4BLFG is found. Such a S4BLFG is identified by angles α_i ($i=0, \dots, 3$) and ϑ_i ($i=1, 2$), which fall into the intervals specified in the previous subsection and summarized in the second column of Tab. (1). The choice of these intervals is somehow arbitrary. For instance, in computing ϑ_2 based on its tangent t_2 , it would have been possible to select ϑ_2 in the interval $]\pi/2, 3\pi/2[$ rather than $]-\pi/2, \pi/2[$. The question also arises on whether, for a given t_2 , both values for ϑ_2 over interval $]-\pi, \pi]$ should be considered in order not to neglect possible S4BLFGs.

Table 2. The set of precision points (deg).

#	φ_1	φ_2
1	0.	0.
2	145.09	–121.30
3	–88.00	–150.31
4	–94.70	78.94
5	75.71	123.08
6	116.10	–174.69

If design parameters α_i ($i=0,...,3$) and ϑ_i ($i=1,2$) were all let free to take values over interval $]-\pi, \pi]$, there would be multiple sets of these parameters that correspond to one root of the seventh-order univariate algebraic equation (Eqn. (47)) and satisfy the compatibility equations (Eqs. (15)). Nevertheless, all these sets correspond to the same S4BLFG. As already stated in Section 2, the reason is traceable to the fact that the same S4BLFG can be identified by more than one set of design parameters. The correspondence between sets of parameters and S4BLFGs becomes one-to-one if the range of some design parameters is suitably restricted. A possible choice of range restriction has been mentioned in Section 2; it is summarized in the third column of Tab. (1).

The back-substitution process described in the previous subsection does not ensure that every design parameter is within the corresponding desired range. It is therefore convenient – though not mandatory – to normalize these parameters, i.e., to replace them by an equivalent set that complies with the intervals listed in the third column of Tab. (1).

The normalization process, which involves parameters α_i ($i=1,2$) and ϑ_i ($i=1,2$) only, is straightforward: if α_i is negative, its sign has to be switched and π added to ϑ_i . Subsequently, it might be necessary to subtract 2π from ϑ_i , in order to make it fall within interval $]-\pi, \pi]$.

5. NUMERICAL EXAMPLE

The procedure explained in the previous sections has been employed to solve the dimensional synthesis of a S4FBLFG that has to comply with the six precision points listed in Tab. (2). The values of angles φ_i ($i=1,2$), expressed in degrees, are reported in Tab. (2) without truncation or round-off errors (i.e., they are exactly expressed by the shown digits).

For the considered set of precision points, the seven roots of Eqn. (47) are all real, positive, and smaller than one. Back substitution yields a S4FBLFG for six of these roots. For the remaining root, the value of u_0 is bigger than one, which means that some design parameters of the corresponding S4FBLFG – and the S4FBLFG itself – are complex.

The six design parameters of the six real S4FBLFGs, expressed in degrees, are reported in Tab. (3) in the normalized form defined by Tab. (1). Each numeric result in Tab. (3),

Table 3. Design parameters of possible function generators (deg).

α_0	50.781238414294	17.271967163613	61.840954313895
α_1	55.310339711106	33.069758962978	63.659497120744
α_2	33.339711881030	15.073096888218	47.741409833391
α_3	70.598420832348	36.133544110469	78.544453432015
θ_1	157.012644242471	146.412865146337	35.547278547488
θ_2	147.701582226003	124.500277290821	-70.138148618823
α_0	27.398660171286	29.780839203761	88.361919930879
α_1	55.701013460318	67.674709378883	61.072883724685
α_2	28.414725048153	34.597029704164	76.862717994887
α_3	55.034544373069	62.131100835442	69.661832055085
θ_1	54.772625720135	80.619961069095	84.255609204148
θ_2	-24.067838148969	18.627280058304	48.483626372627

though necessarily truncated, is endowed with a rather long mantissa in order to make it easier for the reader to check its correctness.

It has been verified that the computed S4FBLFGs – the six real and the one complex – all satisfy the compatibility equations (Eqn. (15)). Therefore no extraneous solution has been introduced by the presented procedure. As already pointed out, despite the six real S4FBLFGs can be assembled in all of the six configurations that correspond to the precision points listed in Tab. (2), further scrutiny is mandatory before selecting any of them as a function generator.

6. CONCLUSIONS

A new procedure to determine all spherical four-bar linkages that comply with six arbitrarily chosen precision points has been presented. Differently from the algorithms available in the literature, the proposed procedure does not introduce extraneous solutions. It results into a seventh-order polynomial equation whose roots are conducive to as many function generators in the complex domain. A numerical example has shown application of the proposed procedure to a case study.

REFERENCES

- [1] Hartenberg, R.S., and Denavit, J., 1964, *Kinematic Synthesis of Linkages*, McGraw-Hill Book Company.
- [2] Cervantes-Sánchez, J.J., Medellín-Castillo, H.I., Rico-Martínez, J.M., and González-Galván, E.J., 2009, "Some improvements on the exact kinematic synthesis of spherical 4R function generators", *Mechanism and Machine Theory*, **44**, pp. 103–121.
- [3] Zimmerman, J.R., 1967, "Four-Precision-Point Synthesis of the Spherical Four-Bar Function Generator", *Journal of Mechanisms*, **2**, pp. 133-139.

- [4] Rao, A.V.M., Sandor, G.N., Kohli, D., and Soni, A.H., 1973, "Closed Form Synthesis of Spatial Function Generating Mechanism for the Maximum Number of Precision Points", *Trans. ASME Journal of Engineering for Industry*, **95**, pp. 725–736.
- [5] Chiang, C.H., 1976, "Synthesis of Spherical Four-Bar Function Generators by Means of Three Relative Poles", *Mechanism and Machine Theory*, **11**, pp. 285-294.
- [6] Lakshminarayana, K., 1972, "On the Synthesis of the Spherical Four-Bar", *Mechanism and Machine Theory*, **7**, pp. 63-69.
- [7] Alizade, R.I. and Kilit, Ö., 2005, "Analytical synthesis of function generating spherical four-bar mechanism for the five precision points", *Mechanism and Machine Theory*, **40**, pp. 863–878.
- [8] Chiang, C.H., 2000, *Kinematics of Spherical Mechanisms*, Krieger Publishing Company.
- [9] Dukkupati, R.V., 2001, *Spatial Mechanisms: Analysis and Synthesis*, CRC Press.
- [10] Cervantes-Sánchez, J.J., Gracia, L., Rico-Martínez, J.M., Medellín-Castillo, H.I., and González-Galván, E.J., 2009, "A novel and efficient kinematic synthesis approach of the spherical 4R function generator for five and six precision points", *Mechanism and Machine Theory*, **44**, pp. 2020-2037.
- [11] Bagci, C., 1984, "Geometric Methods for the Synthesis of Spherical Mechanisms for the Generation of Functions, Paths and Rigid-Body Positions Using Conformal Projections", *Mechanism and Machine Theory*, **19**(1), pp. 113-127.
- [12] Sodhi, R.S., Wilhelm, A.J., and Shoup, T.E., 1985, "Design of a Four-Revolute Spherical Function Generator with Transmission Effectiveness by Curve Matching", *Mechanism and Machine Theory*, **20**(6), pp. 577-585.
- [13] Suixian1, Y., Hong, Y., and Tian, G.Y., 2009, "Optimal selection of precision points for function synthesis of spherical 4R linkage", *Proc. IMechE, Part C*, **223**, pp. 2183-2189.
- [14] Chace, T.R. and Mirth, J.A., "Circuits and Branches of Single-Degree-of-Freedom Planar Linkages", *Trans. ASME Journal of Mechanical Design*, **115**(2), pp. 223-230.

INFLUENZA DELL'ATTRITO SULLO STATO DI TENSIONE E DI DEFORMAZIONE IN ANELLI DI TENUTA IN ELASTOMERO

Edzeario Prati
Dipartimento di Ingegneria Industriale
Università di Parma
E-mail: prati@ied.unipr.it

Marco Silvestri
Dipartimento di Ingegneria Industriale
Università di Parma
E-mail: silve@ied.unipr.it

Alessandro Tasora
Dipartimento di Ingegneria Industriale
Università di Parma
E-mail: tasora@ied.unipr.it

Tito Marin
Dipartimento di Ingegneria Industriale
Università di Parma
E-mail: marin@ied.unipr.it

Sommario. *Questo lavoro presenta alcuni aspetti di rilevante interesse ingegneristico relativi al contatto tra metallo e tenute in elastomero, sia in termini di attrito che di stato tensionale nelle tenute stesse. Sono mostrati due esempi di anelli di tenuta, ciascuno con le sue peculiarità ma anche con alcuni tratti comuni; entrambi sono analizzati con simulazioni ad elementi finiti. Sono poi sviluppate considerazioni circa la stima dell'attrito e come questo influenzi lo stato tensionale e infine sono proposte riflessioni sulle procedure impiegate.*

Parole chiave: *attrito, contatto, tenute, elastomeri*

1. INTRODUZIONE

Da diversi decenni ormai gomme ed elastomeri fanno parte dei materiali a disposizione degli ingegneri per impieghi nei più svariati campi. Questi materiali posseggono tipicamente una struttura amorfa, diversamente da quella cristallina dei metalli e delle ceramiche, e ciò influenza e determina le proprietà meccaniche (resistenza, rigidità, resilienza) e tribologiche, ma anche la loro dipendenza da fattori quali la temperatura e la velocità con cui è impressa una deformazione. Di particolare interesse è l'attrito che si genera nel contatto di un corpo con un altro in elastomero, esempi comuni sono gli pneumatici, le

cinghie, le tenute, ecc.. L'attrito degli elastomeri differisce da quello più noto dei materiali metallici, sia in termini di valori tipici del coefficiente di attrito sia nella natura stessa del fenomeno [1]. Le ragioni di ciò risiedono nuovamente nella microstruttura del materiale (lunghe catene polimeriche intrecciate) da cui segue una scarsa rigidità e un elevato attrito interno [2]. Quest'ultimo d'altro canto determina l'apprezzata capacità di agire come smorzatore di vibrazioni o di dissipare energia dovuta ad impatti. La sostanziale differenza tra l'attrito di un metallo e di un elastomero è da imputarsi al tipo di deformazione locale nei punti di contatto. Nel primo caso infatti tale deformazione è di tipo plastico, mentre nel secondo essa è di natura puramente elastica con l'elastomero che, essendo molto cedevole, si adatta alle asperità della superficie dell'altro corpo con un effetto netto di un contatto locale molto più esteso e quindi un coefficiente di attrito tipicamente più elevato. In presenza di olio interposto tra i corpi, diversi regimi di lubrificazione, con caratteristiche assai diverse, sono poi raggiungibili e ciò modifica ulteriormente lo scenario e il concetto stesso di attrito [3].

In questo lavoro vengono considerati due casi che pur essendo apparentemente piuttosto diversi tra loro presentano in realtà alcuni tratti comuni. Entrambi sono affrontati con simulazioni ad elementi finiti. Il primo riguarda una tenuta radiale a labbro per alberi rotanti in presenza di olio lubrificante, il secondo riguarda invece una tenuta a sezione quadrata, quindi con ampia superficie di contatto, nel caso di spostamento assiale dell'albero. Vengono discussi in entrambi i casi la deformazione e lo stato tensionale nelle tenute così come le tensioni (normali e tangenziali) che si sviluppano nelle zone di contatto. Considerazioni riguardo all'attrito sono presentate nelle varie sezioni fino ad una breve discussione finale nelle conclusioni.

2. TENUTA RADIALE A LABBRO

In questa Sezione si presentano alcuni risultati relativi ad una tenuta radiale a labbro avente dimensioni 70x110x12 mm in elastomero a base nitrilica (NBR). Si tratta di una tenuta commerciale, destinata ad impieghi con alberi rotanti e in condizioni di lubrificazione, generalmente con oli minerali.

Lo studio si è focalizzato sulla distribuzione della pressione di contatto, nell'interfaccia albero-tenuta, dovuta all'interferenza di montaggio e alla presenza di olio in pressione. Il modello ad elementi finiti ha sfruttato l'assialsimmetria della geometria e del carico perciò la mesh si è ridotta ad una sezione piana della tenuta modellata con elementi assialsimmetrici. La geometria della tenuta è stata ricostruita basandosi su immagini digitali ad elevato ingrandimento di una tenuta (nuova e non usurata) sezionata radialmente. La porzione della tenuta in NBR è stata modellata come materiale iperelastico e incompressibile, la parte metallica come lineare elastico, mentre l'albero è stato semplificato con una superficie rigida. La molla è schematizzata con un materiale fittizio avente rigidità equivalente. Non è specificato alcun attrito all'interfaccia albero-tenuta così come non è considerato la presenza di lubrificante. Maggiori dettagli sulla procedura sono reperibili in [4].

In Fig. 1 è mostrata la deformata della tenuta in tre diverse configurazioni: sola interferenza di montaggio (pari a 1.5 mm sul diametro), interferenza più pressione dell'olio pari a 0.3 MPa e 0.6 MPa. Il lamierino metallico, inglobato nel corpo dell'anello di tenuta e con funzione di irrigidimento della tenuta stessa, non è mostrato per semplicità.

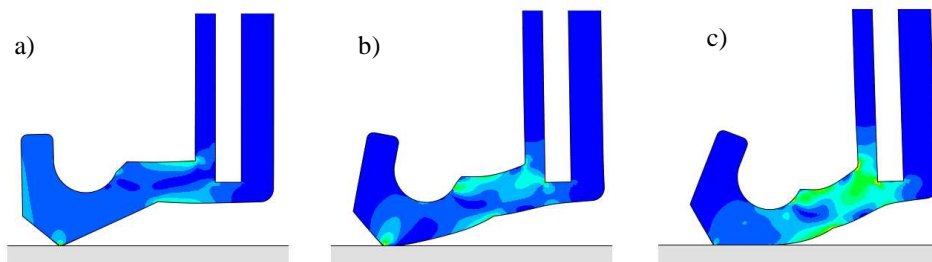


Figura 1. Modello ad elementi finiti della tenuta radiale, a) interferenza di montaggio, b) pressione dell'olio pari a 0.3 MPa, c) pressione dell'olio pari a 0.6 MPa. Scala della deformata $\times 1$, la mappa riporta la distribuzione della tensione di V.Mises.

Concentrazioni di tensione sono evidenti nella zona di contatto e in quelle con elevata flessione. Ad elevata pressione dell'olio si nota la tendenza della tenuta a schiacciarsi sull'albero aumentando l'area di contatto. Quando questa deformazione diventa eccessiva si ha perdita di funzionalità della tenuta (dal catalogo della tenuta in esame si scopre un limite pari a 0.5 MPa, ben giustificato dall'eccessiva deformazione di Fig. 1c).

La distribuzione della pressione di contatto all'interfaccia albero-tenuta è proposta in forma grafica in Fig. 2 per alcuni valori di pressione nel lato olio. L'ascissa riporta la coordinata assiale (z) la quale, per favorire la leggibilità dei risultati, è stata traslata, per ciascun valore di pressione, in modo da avere lo zero sempre in corrispondenza del primo nodo della tenuta che è andato a contatto con l'albero. L'interpretazione di queste curve è, alla luce delle deformate in Fig. 1, evidente: la pressione dell'olio influenza fortemente la geometria della tenuta nel suo insieme, ma anche la zona di contatto locale con l'albero. Questo grafico può essere rielaborato per estrarre in forma condensata alcune grandezze che ben sintetizzano l'evoluzione del contatto. Ciò è presentato in Fig. 3 dove la larghezza di contatto (b) la pressione media (p_m) e quella massima (p_{max}) di contatto sono plottate nel range di pressione nel lato olio qui considerato. La pressione massima è il picco di tensione normale all'interfaccia, riscontrato sull'intera porzione di tenuta effettivamente a contatto con l'albero. Si nota che questo valore prima cresce nell'intervallo di pressione dell'olio 0-0.3 MPa per poi diminuire. Anche la pressione media (p_m) ha un andamento parzialmente simile anche se il picco è raggiunto quando la pressione dell'olio è prossimo a 0.1 MPa. Dopo circa 0.35 MPa di pressione dell'olio si vede che la pressione media scende sotto al valore che si ha con sola interferenza di montaggio. L'unico parametro che è crescente, in modo nonlineare, con la pressione dell'olio è la larghezza di contatto (b). Questa passa da circa 70 μm con sola interferenza di montaggio a circa 2.2 mm con massima pressione dell'olio. La diminuzione della pressione media si spiega quindi con l'incremento dell'area di contatto ossia la forza totale scambiata si distribuisce progressivamente su una superficie maggiore. La pressione di contatto integrata su tutta la superficie cilindrica di contatto determina la forza radiale F_r complessiva esercitata dall'olio, tramite la tenuta, sull'albero. L'andamento di F_r , riportato in Fig. 3b, è monotono e crescente in modo circa lineare con la pressione dell'olio.

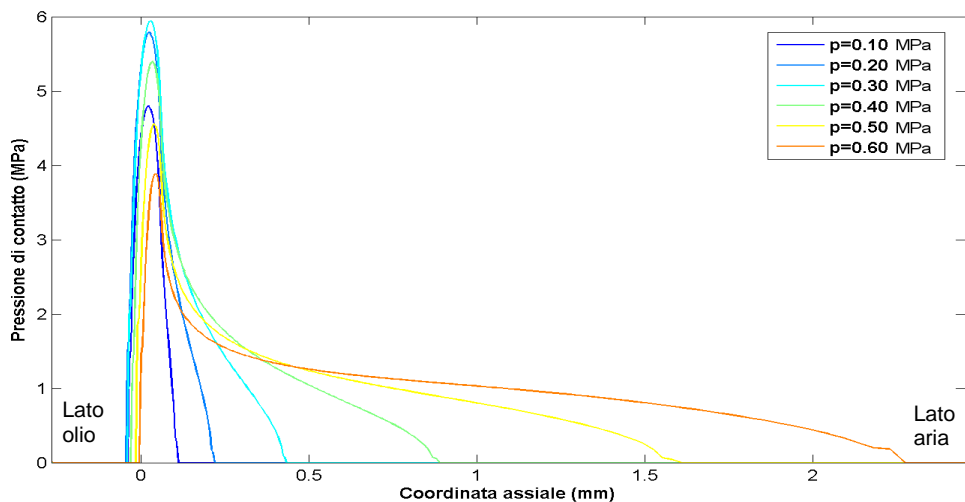


Figura 2. Distribuzione della pressione di contatto, in direzione assiale, al variare della pressione nel lato olio (p).

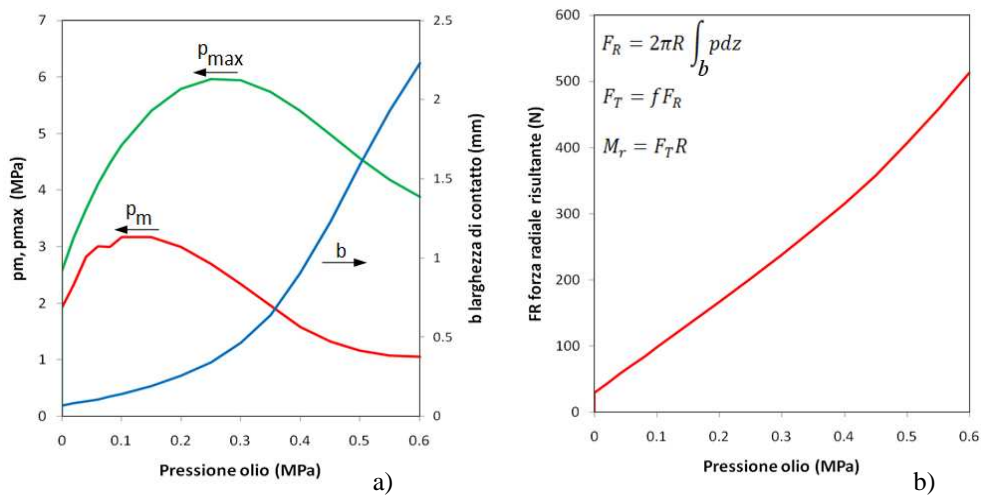


Figura 3. a) Pressione massima, media e larghezza di contatto in funzione della pressione dell'olio; b) forza radiale risultante in funzione della pressione nel lato olio.

Data la forza radiale risultante (F_r) calcolata con questa analisi statica e dato un coefficiente di attrito (f), si potrebbe stimare una forza tangenziale ($F_t = F_r \cdot f$) e da questa un momento resistente ($M_r = F_t \cdot R$, con R raggio dell'albero) interpretabile in primissima approssimazione come un momento resistente di primo distacco.

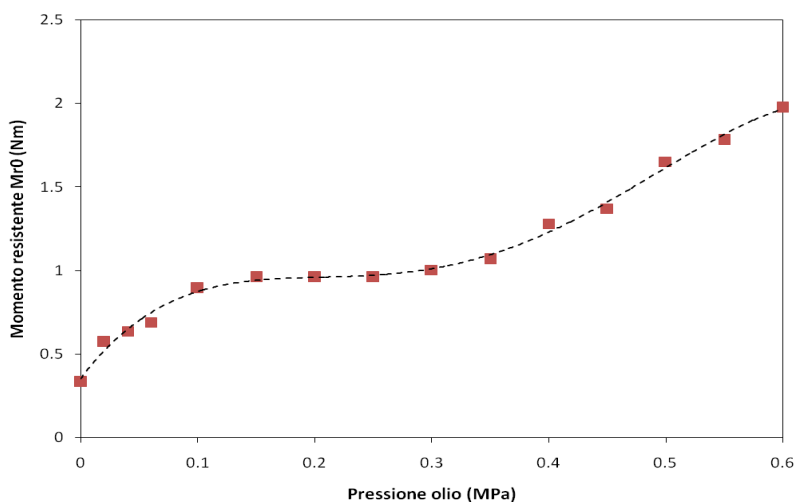


Figura 4. Tipico andamento, misurato sperimentalmente, della coppia di spunto in un albero con tenuta radiale a labbro e presenza di olio lubrificante. La linea tratteggiata è una polinomiale di quarto grado che ben rappresenta l'andamento sperimentale.

Sperimentalmente si osserva che l'andamento della coppia di spunto, per questa tenuta e in condizioni di temperatura ambiente, è quello mostrato in Fig. 4 [4]. Questa curva ha sì un andamento monotono crescente, ma non è sovrapponibile a quello di F_r in Fig. 3b. Si evince da ciò che il coefficiente di attrito non può essere assunto costante ma che esso deve essere funzione dei diversi parametri in gioco tra i quali, se si escludono per semplicità i fenomeni legati alla lubrificazione all'interfaccia, principalmente la pressione locale di contatto. In sostanza, relazioni fenomenologiche, tipo la legge di Coulomb, non sono applicabili in casi come quello discusso in questa Sezione. Ciò ci riporta a quanto detto nell'introduzione, il contatto in presenza di uno o più corpi deformabili ha un comportamento diverso da quello più semplice e più noto di contatto tra corpi rigidi (metallici). Si possono indicare almeno tre fattori che influenzano l'attrito (secco) negli elastomeri: la pressione, la temperatura e la velocità di scorrimento relativa tra le superfici. Nel caso più complesso di presenza di lubrificante occorre aggiungere alla lista anche la viscosità.

Rimandando alle letterature [1] per una più dettagliata descrizione, si ricorda qui solo un aspetto notevole dell'attrito negli elastomeri, cioè la variazione del coefficiente di attrito con la pressione di contatto. Questa variazione si manifesta solitamente con una riduzione del coefficiente di attrito con la pressione $f(p)$, come documentato fin dai classici lavori di Thirion [5] e Schallamach [6]. In letteratura la dipendenza $f(p)$ è riscontrabile come leggi empiriche o semi-empiriche e la sperimentazione diretta rimane per lo più lo strumento principale per caratterizzarla.

Alla luce delle considerazioni appena espresse, si è applicata una di queste relazioni tratta da [7] per rielaborare i risultati dell'analisi ad elementi finiti della tenuta radiale. La procedura, che è stata ripetuta per ogni condizione di carico (cioè per i diversi valori della pressione dell'olio), è molto semplice, e per questo appetibile dal punto di vista del tecnico che cerca metodi sicuri e di immediata applicabilità, e consiste nei passaggi di seguito

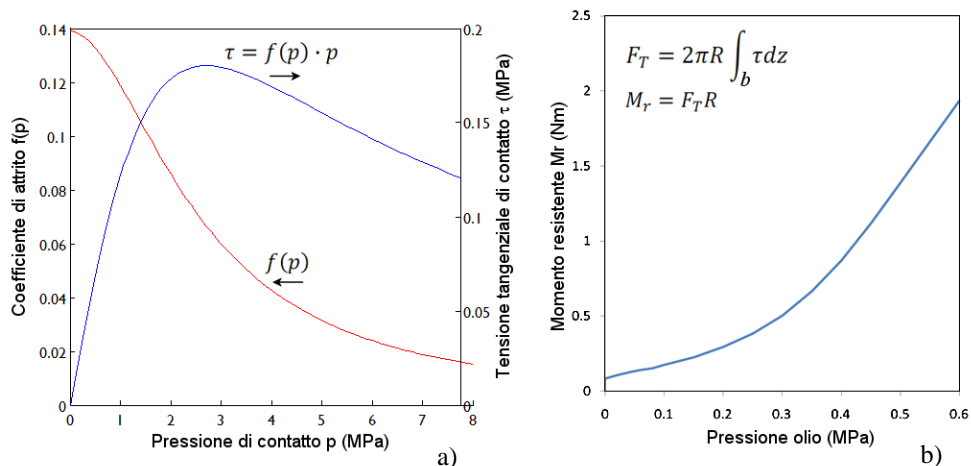


Figura 5. a) Coefficiente di attrito [7] e tensione tangenziale in funzione della pressione di contatto, b) momento resistente al variare della pressione dell'olio.

descritti. In primo luogo, data la distribuzione di pressione di contatto (p) si è calcolata una tensione tangenziale (τ), nodo per nodo, come il prodotto della pressione per il coefficiente di attrito locale (cioè dipendente dalla pressione di contatto in quel punto). Successivamente si è integrata la distribuzione di pressione tangenziale su tutta la superficie di contatto tenuta-albero per ottenere la forza tangenziale risultante (F_t). Infine da questa si è ottenuta la coppia resistente moltiplicandola per il raggio dell'albero (R). La rappresentazione grafica di quanto esposto è riportata in Fig. 5.

Questa procedura ha diverse criticità: la relazione $f(p)$ tratta dalla letteratura era stata ricavata per una geometria molto diversa e in condizioni di assenza di lubrificante; l'attrito è introdotto solo dopo che la distribuzione di pressione normale era stata ottenuta con una simulazione senza attrito (la deformata della tenuta non è perciò influenzata dall'attrito tenuta-albero); l'effetto del velo di lubrificante all'interfaccia è totalmente escluso a priori dal calcolo. Ciononostante si è ritenuto che fosse comunque istruttivo verificare se ciò avrebbe aiutato ad ottenere una miglior correlazione con i risultati sperimentali. Dal confronto di Fig. 5b con Fig. 4 si evidenzia invece che, nonostante gli ordini di grandezza siano decisamente comparabili e che i valori della coppia resistente siano verosimili, questa procedura non porta alcun vantaggio riguardo all'andamento della curva M_r in funzione della pressione dell'olio.

Lo studio delle tenute radiali a labbro con presenza di olio e in condizioni di spunto non è quindi affrontabile appieno con le procedure sopra esposte (coefficiente di attrito costante, coefficiente di attrito in funzione della pressione locale di contatto) ma richiedono strumenti più sofisticati. Si ritiene inoltre che non si possa prescindere dal film d'olio all'interfaccia albero-tenuta e perciò la metodologia di simulazione deve tener conto sia della deformabilità della tenuta che dell'effetto di sostentamento possibilmente indotto dal lubrificante. In un precedente lavoro [8] gli autori hanno però mostrato come, con albero

rotante e in condizioni di regime, con un metodo simile basato su una simulazione ad elementi finiti, sia possibile studiare la coppia di attrito. In [8] era stato inoltre proposto una

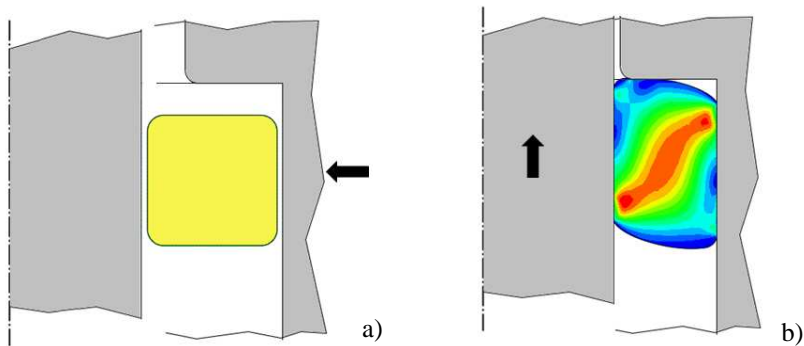


Figura 6. a) Fase di montaggio con interferenza della tenuta, simulata mediante uno spostamento radiale della sede; b) scorrimento assiale dello stelo ed eventuale trascinamento della tenuta, per attrito, contro la battuta laterale.

procedura sperimentale per la misura della coppia di attrito in presenza di eccentricità dinamica.

3. TENUTA A SEZIONE QUADRATA

L'esempio proposto concerne una tenuta di sezione quadrata (con vertici arrotondati) interposta tra uno stelo cilindrico e una sede. La simulazione ha riguardato due condizioni, la prima data dal solo montaggio con interferenza, la seconda da un movimento assiale dello stelo. Il sistema assialsimmetrico è schematizzato in Fig. 6. Il materiale elastomerico è modellato come iperelastico con le medesime proprietà di quello usato nella Sezione precedente. Lo stelo ha un diametro di 70 mm, la sezione della tenuta ha una dimensione radiale pari a 2 mm in condizioni indeformate la quale passa a 1.6 mm in seguito al montaggio con interferenza (la deformazione radiale nominale è perciò del 20%). Stelo e sede sono state modellate come superfici rigide. All'interfaccia tra la tenuta e le parti metalliche è stato definito un contatto con attrito che, per semplicità, segue una legge di tipo coulombiano con coefficiente costante. Diversi valori del coefficiente di attrito sono stati usati, qui si riportano i risultati relativi a tre casi: $f=0.1$, 0.3 , 0.5 .

La distribuzione di pressione e di tensione tangenziale nelle due zone di contatto (stelo-tenuta, tenuta-sede) al termine della fase di montaggio sono mostrate in Fig. 7b. L'ascissa del grafico corrisponde alla posizione dei nodi della mesh lungo il perimetro della tenuta. Per quanto riguarda la pressione sia la distribuzione in A che quella in B sono simmetriche come ci si attende. Le curve in B sono leggermente più basse di quelle in A poiché, a parità di forza radiale risultante scambiata, il contatto in B avviene su una superficie cilindrica più ampia. Gli andamenti della pressione sono uguali per $f=0.3$ e $f=0.5$ mentre per il caso $f=0.1$ le curve hanno valori inferiori ma ciò è compensato dall'incremento nell'ampiezza del contatto. Vi sono picchi alle estremità delle zone di contatto con valori circa 1.5 più elevati rispetto alla zona centrale, ciò si rispecchia anche nella distribuzione di tensione come

mostrato in Fig. 7a. La tensione tangenziale ha andamenti antisimmetrici (con risultante netta nulla) e mostra una più marcata differenziazione tra i casi $f=0.3-0.5$ e il caso $f=0.1$.

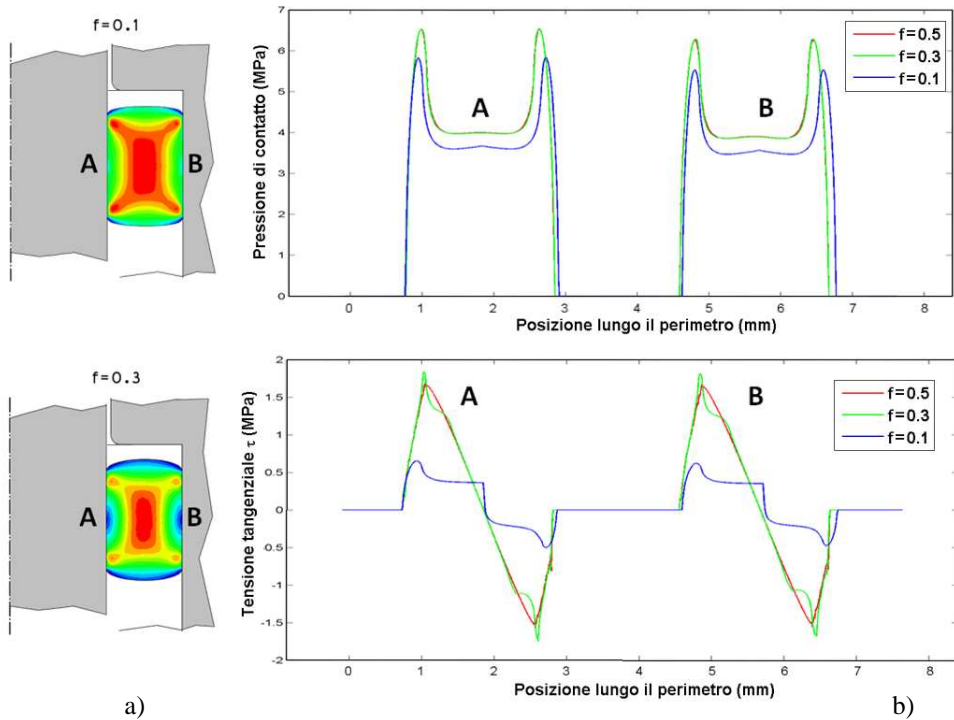


Figura 7. Fase di montaggio per interferenza della tenuta all'interno della propria sede; a) deformata della tenuta e tensione di Von Mises; b) pressione normale e tensione tangenziale di contatto lungo il perimetro della tenuta.

Per elevati valori di f l'andamento è tipo onda triangolare, per bassi valori è tipo onda quadra. Pressione normale e tensione tangenziale modificano radicalmente gli andamenti quando viene applicata una traslazione assiale dello stelo. Vi è una fase iniziale nella quale la tensione tangenziale all'interfaccia rimane mediamente inferiore a quella critica, ossia a quella necessaria ad indurre il moto relativo (che nel caso di attrito coulombiano è proporzionale, punto per punto, alla pressione locale di contatto moltiplicata per il coefficiente di attrito) e si ha perciò trascinamento della tenuta fino a che questa non tocca, nel punto C, la battuta della sede. In tale condizione, per $f=0.1$ e $f=0.3$ la tensione tangenziale in A supera quella critica e si ha scorrimento dello stelo. Per $f=0.5$, invece, ciò avviene solo dopo che lo stelo si è trascinato parte della tenuta nell'interstizio stelo-sede, a questo punto però la mesh si distorce troppo e la soluzione non raggiunge più convergenza. La Fig. 8 riporta le condizioni di tensione di Von Mises, pressione normale e tensione tangenziale nei contatti a seguito di uno scorrimento dello stelo pari a 0.75 mm. Questo rappresenta circa il momento in cui, per il caso $f=0.1$, si abbandona il trascinamento ed inizia lo scorrimento dello stelo sulla tenuta nel punto A. La zona maggiormente sollecitata

è una striscia di materiale che percorre circa diagonalmente la sezione, la massima tensione di Von Mises, così come la massima deformazione principale, si hanno per $f=0.5$, quelle

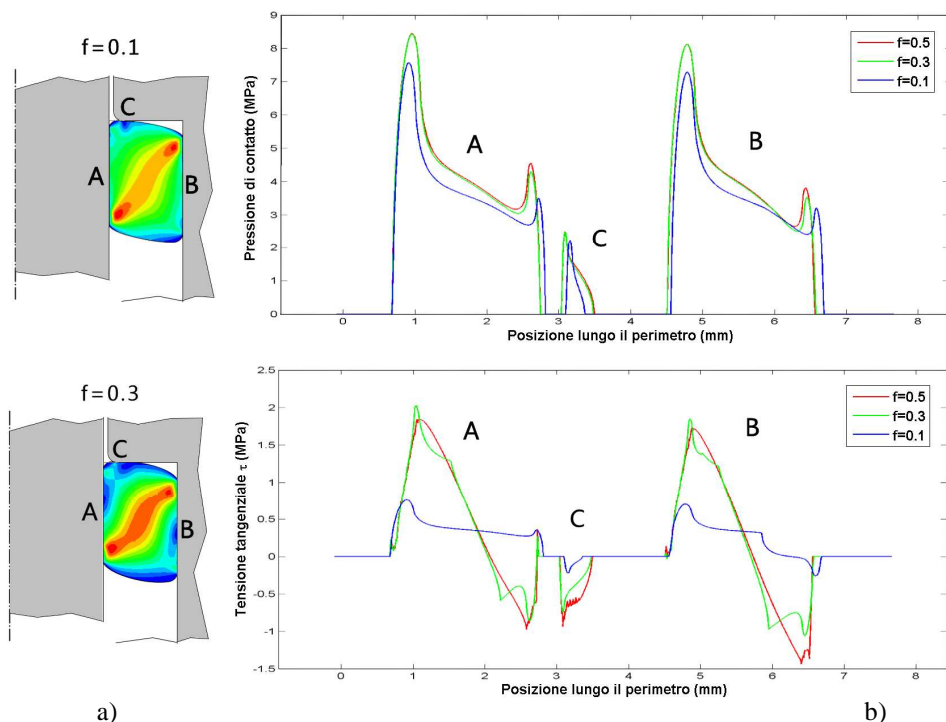


Figura 8. Fase di scorrimento assiale in direzione verticale dello stelo; a) deformata della tenuta e tensione di Von Mises; b) pressione normale e tensione tangenziale di contatto lungo il perimetro della tenuta.

minime per $f=0.1$. Si nota come la diagonale di materiale più sollecitato parta e termini circa in corrispondenza dei due picchi di pressione normale e tensione tangenziale di contatto nelle zone A e B. Il movimento assiale dello stelo distorce gli andamenti della pressione normale (ora non più simmetrici) e di quella tangenziale (ora non più antisimmetrici). Inoltre la risultante della tensione tangenziale di contatto (in direzione assiale) è ora diversa da zero ed pari alla forza assiale necessaria a muovere lo stelo.

4. DISCUSSIONE E CONCLUSIONI

I due casi descritti hanno evidenziato alcune delle problematiche tipiche nel contatto tra metallo e tenute in elastomero, in presenza di lubrificante e non. Nel primo esempio si è discussa la dipendenza del coefficiente di attrito da svariati parametri tra cui la pressione locale di contatto. L'applicazione di una relazione $f(p)$ empirica non ha migliorato la correlazione tra il momento resistente calcolato e la coppia di spunto misurata sperimentalmente. Si è inoltre

visto come all'aumentare della pressione dell'olio il contatto tra tenuta ed albero passi da alcune decine di micrometri a circa 2 mm. In queste condizioni di contatto esteso si ha una somiglianza con il secondo caso, quello della tenuta a sezione quadra. In quest'ultimo esempio poi il contatto è stato studiato, oltre che in condizioni di sola interferenza stelo-tenuta, anche in condizioni di spostamento assiale dello stelo, con forte deformazione indotta nell'elastomero. Per bassi coefficienti di attrito si determina l'eventuale perdita di aderenza e lo scorrimento dello stelo sulla tenuta. Si è mostrato come il valore del coefficiente di attrito influenzi, sia la distribuzione delle pressione normale e della tensione tangenziale di contatto, che il campo tensionale nella tenuta. Entrambi i casi sono stati analizzati con simulazioni statiche col metodo degli elementi finiti. Diversi limiti sono emersi e sono principalmente legati alla definizione delle condizioni di attrito all'interfaccia.

Sebbene sia nella sua fase iniziale, questo studio presenta diverse considerazioni di base per la comprensione dei fenomeni di deformazione e attrito nel contatto tra elastomero e metallo, in particolare per la comunissima applicazione delle tenute radiali a labbro con elevate pressioni dell'olio.

BIBLIOGRAFIA

- [1] Smith, R.H., 2008. Analyzing friction in the design of rubber products and their paired surfaces. CRC Press.
- [2] Besdo, D., Heimann, B., Klüppel, M., Kröger, M., Wriggers, P., Nackenhorst U., 2010. Elastomere Friction. Springer.
- [3] Hamrock, B.J., Schmid, S.R., Jacobson, B.O., 2004. Fundamental of fluid film lubrication (2nd Ed). Marcel Dekker.
- [4] Silvestri, M., Prati, E., Tasora, A., Marin, T., 2010. "Frictional behaviour of radial lip seals on varying exerted pressure". Proceedings of the NordTrib 2010 Conference, Lulea Sweden.
- [5] Thirion, P., 1948. "Les coefficients d'adhérence du caoutchouc", *Rubber Chemistry and Technology*, **21**, 505-15.
- [6] Schallamach, A., 1958. "Friction and abrasion of rubber", *Wear*, **1**, 384-417.
- [7] Jimenez, M.A., Bielsa, J.M., Rodriguez, R., Dobon, S., 2006. "The influence of contact pressure on the dynamic friction coefficient in cylindrical rubber-metal contact geometries". Proceedings of the IUTAM Symposium on Computational Methods in Contact Mechanics, Hannover Germany.
- [8] Silvestri, M., Prati, E., Tasora, A., 2006. "Radial lip seals efficiency under dynamic operating conditions", Proceedings of the AITC-AIT 2006 International Conference on Tribology, Parma Italy.

METODI DI PROGETTAZIONE INTEGRATA DI SISTEMI AUTOMATICI ADATTIVI

Angelo O. Andrisano, Marco Faretra, Francesco Gherardini, Alessandro Guerra, Francesco Leali, Marcello Pellicciari, Fabio Pini, Alberto Vergnano

Dipartimento di Ingegneria Meccanica e Civile,

Università di Modena e Reggio Emilia

E-mail: angelo.andrisano@unimore.it

Sommario. *Le prestazioni dei moderni sistemi automatici di assemblaggio derivano sempre più dall'integrazione di apparati meccanici con attuatori programmabili, sensori, sistemi di elaborazione dati e complesse logiche di controllo. La disponibilità di tali tecnologie permette di affrontare nuove sfide tecniche ma comporta una rivoluzione nel modo di concepire meccanismi e macchine. I tradizionali modelli di progettazione, orientati alla definizione per funzioni degli apparati e strutturati con l'intervento sequenziale delle diverse aree di conoscenza dell'ingegneria, non permettono di sfruttare tutte le potenzialità delle diverse tecnologie coinvolte e di progettare comportamenti adattivi. È dunque necessario costruire nuovi metodi di progettazione che consentano di produrre sistemi in grado di interagire con l'ambiente circostante, di adattare autonomamente il proprio comportamento e di massimizzare e/o ottimizzare le prestazioni.*

Il presente lavoro descrive un metodo per la progettazione integrata di sistemi automatici adattivi attraverso la modellazione e la prototipazione virtuale dei comportamenti. Il metodo segue un approccio sistematico e considera, a diversi livelli di dettaglio, l'evoluzione dinamica dei comportamenti sia tempo-continua che guidata da eventi discreti. L'integrazione delle conoscenze e delle tecnologie in un ambiente virtuale può portare all'evoluzione in nuovi principi risolutivi, alla creazione delle basi di conoscenza su cui sono basate le logiche adattive, ad avanzate strategie di sensorizzazione ed attuazione, all'ottimizzazione complessiva del sistema ed alla compressione dei tempi di sviluppo.

Questa memoria riporta infine l'applicazione del metodo proposto nella progettazione di un sistema di assemblaggio automatico.

Parole chiave: *progettazione integrata, sistemi automatici, comportamenti adattivi, prototipazione virtuale*

1. INTRODUZIONE

Le tecnologie oggi disponibili nel settore dell'automazione industriale consentono l'implementazione di complessi modelli di ottimizzazione del comportamento delle macchine e di interazione con l'ambiente operativo attraverso opportune strategie di sensorizzazione ed attuazione. L'evoluzione di queste tecnologie abilita spesso innovazioni radicali tanto nelle operazioni eseguite quanto nelle stesse architetture funzionali delle macchine, progettate non solo per essere flessibili, ma anche per offrire una reale adattività dei comportamenti alle variazioni delle condizioni operative.

La flessibilità è definita in letteratura come la capacità di un sistema di modificare alcune variabili o assumere diversi stati in risposta a variazioni dei requisiti con una

contenuta perdita in termini di tempo, costi o prestazioni, [1, 2]. Questa definizione non esclude la flessibilità o la riconfigurabilità a seguito di intervento manuale degli utilizzatori, anche a sistema fermo. Un sistema adattivo è invece un sistema capace di modificare autonomamente il proprio comportamento per perseguire i propri obiettivi anche in presenza di variazioni delle condizioni operative. La letteratura di riferimento tratta ampiamente i diversi aspetti e declinazioni dell'adattività dei sistemi. Se da un lato questo tema è ampiamente sviluppato nell'ingegneria del software, l'ingegneria meccanica incorpora l'adattività in un approccio più ampio, consistente nella progettazione di sistemi capaci di adattarsi a diversi tipi di incertezze, [3, 4]. I sistemi adattivi possono quindi essere considerati un sottoinsieme dei sistemi flessibili. Per i sistemi adattivi in senso stretto inoltre si fa spesso riferimento al concetto di adattività rispetto alle variazioni prevedibili o (in una certa misura) imprevedibili dei parametri di lavoro interni ed esterni al sistema stesso, aspetto questo non propriamente considerato dai sistemi flessibili.

L'architettura dei sistemi produttivi adattivi è generalmente concepita in modo modulare, sia dal punto di vista meccanico che della logica di comportamento, aspetto, quest'ultimo, di grande interesse scientifico. I sistemi adattivi più avanzati, infatti, sono composti di agenti isolati ma interagenti in una rete di comunicazione, mentre realizzano il comportamento adattivo autonomamente o in modo parzialmente guidato dai livelli di controllo gerarchicamente superiori. Tale comportamento adattivo si traduce, generalmente, in tre diverse azioni [5]: osservazione delle variabili interne e dell'ambiente esterno, ottimizzazione degli obiettivi in relazione a certe leggi di risposta e adattamento dell'azione di risposta per conseguire l'obiettivo ottimizzato.

I comportamenti adattivi possono variare dalla modifica online dei parametri, [6], fino all'adattività strutturale, [7]. In letteratura, così come nella realtà industriale, l'adattività è solitamente guidata dall'ingegneria del software, [8, 9], che deve pertanto essere integrata nelle attività di progettazione meccanica per sviluppare ed ottimizzare i comportamenti, risultanti dall'integrazione degli apparati meccanici ed elettronici e del software.

2. APPROCCIO TRADIZIONALE ALLA PROGETTAZIONE

I flussi di progettazione possono essere suddivisi a seconda della direzione di progettazione in Top-Down e Bottom-Up: dal generale al particolare, o viceversa. In letteratura si ritiene più promettente il primo, [10], anche se alcune fasi di progettazione dovranno per forza essere eseguite con un flusso Bottom-Up, come per l'introduzione di componentistica e moduli mecatronici programmabili o nel riutilizzo di blocchi funzionali software. I principali vantaggi di questo approccio alla progettazione sono l'individuazione anticipata di errori e punti deboli del sistema, la risoluzione anticipata dei requisiti in conflitto, l'orientamento a minimizzare le incertezze ed i tempi di sviluppo contrapposto alla pratica del trial and error, la facilitazione di redazione di una documentazione esaustiva dell'architettura del sistema da progettare.

Tra i possibili metodi di progettazione Top-Down si può considerare come riferimento, per efficacia e diffusione, un metodo sistematico articolato nelle seguenti quattro fasi, [10]:

1. **Formalizzazione delle specifiche:** analisi e definizione delle proprietà da conseguire con l'attività di progettazione del sistema con l'obiettivo di redigere una lista di requisiti, o specifiche di progetto.
2. **Progettazione concettuale:** scomposizione funzionale delle specifiche in problemi, per i quali si ricercano, si valutano e si selezionano i possibili principi risolutivi.

3. **Progettazione esecutiva:** sviluppo dei principi risolutivi in soluzioni tecniche opportunamente dimensionate, verifica delle prestazioni e dell'effettiva realizzabilità tecnica ed economica, ottimizzazione dell'intero sistema.
4. **Progettazione di dettaglio:** definizione di tutte le caratteristiche necessarie per la realizzazione fisica del sistema, redazione dei disegni e dei documenti di produzione.

Questo metodo di progettazione sistematica si sviluppa in fasi successive di analisi e sintesi attraverso la scomposizione funzionale delle specifiche e la progettazione degli apparati necessari per assolvere le diverse funzioni, per giungere, infine, all'integrazione di tali apparati nel sistema complessivo. Questo approccio non tiene efficacemente conto dell'evoluzione del sistema nel dominio del tempo poiché mira spesso a determinare il miglior compromesso tra effetti contrapposti in condizioni nominali (costi/prestazioni, rigidità/massa, lead time /dimensioni magazzino, etc.). Nel caso di sistemi dinamici, però, il comportamento può essere descritto matematicamente con leggi fortemente nonlineari sovrapposte a modellazioni ad eventi discreti, [11], che possono integrarsi, compensarsi o contrapporsi con effetti ben diversi dal comportamento nominale considerato in sede di scomposizione funzionale, specialmente in presenza di variazioni delle condizioni operative. Questo approccio di progettazione di apparati per funzioni deve essere quindi superato con una progettazione di comportamenti nel dominio del tempo, da valutare mediante simulazioni di modelli comportamentali.

La progettazione dei sistemi automatici è un processo molto complesso in cui spesso si rilevano inefficienze nell'integrazione tra le diverse aree disciplinari dell'ingegneria. Gli ingegneri meccanici, responsabili del know-how di prodotto e processo, guidano le fasi iniziali dello sviluppo, definendo rigidamente le soluzioni tecniche. Si progettano quindi gli schemi elettrici e l'hardware per il controllo degli apparati, che vengono poi fisicamente costruiti ed installati. A questo punto possono intervenire i programmatori per sviluppare il software per il controllo dei comportamenti nominali e adattivi, i cicli di funzionamento, nonché tutte le procedure di avviamento, di arresto e di gestione delle situazioni anomale. Le reali logiche di controllo vengono quindi spesso progettate in dettaglio solo durante la messa in servizio finale, [12, 13], mentre i comportamenti vengono ideati a monte dall'ingegneria meccanica. La pratica comune, inoltre, dimostra spesso inefficienze nella comunicazione di informazioni tra l'ingegneria meccanica e quella del software, che porta a ritardi di tempo e a comportamenti non ottimizzati e l'ingegneria del software ha spesso ridotte possibilità di introdurre cambiamenti e migliorie al sistema complessivo, [14, 15].

Occorre, in definitiva, passare da una progettazione di apparati per funzioni in configurazioni statiche e con una scarsa integrazione tra le diverse tecnologie coinvolte ad un metodo basato sulla progettazione integrata di comportamenti adattivi e cicli di lavorazione e sulla progettazione e valutazione tramite modelli virtuali nel dominio del tempo.

3. METODO DI PROGETTAZIONE MEDIANTE PROTOTIPAZIONE VIRTUALE

In questa sezione si descrive un metodo di progettazione basato sull'utilizzo della prototipazione virtuale nelle diverse fasi di sviluppo Top-Down di un sistema di assemblaggio automatico.

Formalizzazione delle specifiche

Nella formalizzazione delle specifiche per un sistema adattivo di assemblaggio automatico occorre considerare che non tutto il sistema sarà adattivo ma solamente alcune sue parti e non sarà inoltre possibile monitorare qualunque grandezza. Vi saranno poi parametri prestazionali da ottenere indipendentemente dalle variazioni dell'ambiente operativo mentre per altri sarà accettabile una maggiore flessibilità, [8, 16].

In questa fase occorre analizzare il dominio di conoscenza utilizzato per riferirsi ad un determinato contesto operativo, [17], raccogliendo modelli e descrizioni nei diversi standard di formalizzazione utilizzati dalle discipline ingegneristiche. L'esperienza dei progettisti di prodotto e di processo, arricchita da simulazioni virtuali CAE, definisce un insieme di dati e regole per le ontologie su cui fondare lo sviluppo dei comportamenti del sistema. Grazie a prototipi virtuali CAE avanzati è possibile prevedere le variazioni delle prestazioni dei prodotti realizzati ammissibili per l'utente finale. Risultano inoltre fondamentali in questa fase le informazioni sul numero di operazioni da eseguire senza variazioni di set-up, sul mix di prodotto all'interno di una piattaforma, sulla variabilità della domanda di mercato e sul tempo accettabile per la consegna del prodotto ed infine sul livello di automatizzazione richiesto per l'attuazione dei comportamenti adattivi.

La formalizzazione delle specifiche per i sistemi adattivi deve intrinsecamente trattare l'incertezza, l'incompletezza, la variabilità e la stocasticità, spostando il livello del discorso da prescrittivo a flessibile. In particolare non tutte le specifiche sono esprimibili quantitativamente sin dalle prime fasi e vengono rifinite man mano che si procede con le fasi di progettazione, mentre la stessa lista delle specifiche costituisce un importante strumento di validazione e verifica dei modelli virtuali che supporterà tutta l'attività di progettazione tramite prototipi virtuali.

Progettazione concettuale

La progettazione concettuale inizia con l'identificazione astratta dei problemi da risolvere, stabilendo un'architettura per i comportamenti del sistema. In particolare i comportamenti adattivi possono essere descritti secondo quattro dimensioni di modellazione, [18]: compiti e risultati attesi dal sistema (*obiettivi del sistema*), leggi e range di variabilità a cui il sistema deve rispondere (*cause di adattività*), modalità di attuazione di una risposta adattiva (*meccanismi per l'adattività*) ed efficacia e velocità di risposta attese e flessibilità consentita per il meccanismo adattivo (*effetti sul sistema*).

Si devono poi individuare le soluzioni tecniche concettuali preposte al conseguimento dei comportamenti adattivi. Tali soluzioni, definite principi risolutivi, devono essere combinate per conseguire i compiti del sistema complessivo. Le diverse possibili strutture organizzative devono essere rapidamente valutate per eliminare scelte tecniche chiaramente non promettenti per gli scopi, i costi o i tempi utili per l'iniziativa aziendale.

Superata la prima fase di scrematura si procede alla prototipazione virtuale delle soluzioni concettuali con un approccio integrato, [19]. Nel caso specifico di sistemi automatici di assemblaggio la modellazione comportamentale risulta più difficile, dovendo trattare variabili logiche e numeriche di diversi formati (booleane, stringhe, reali, intere, etc.).

Nella fase di progettazione concettuale è spesso sufficiente realizzare modelli cinematici dei meccanismi, per quanto riguarda la meccanica, ed eseguire blocchi funzionali solamente per il comportamento nominale per quanto riguarda il software, in quanto in fase di progettazione concettuale non è necessaria una modellazione dettagliata. Possono anche essere omessi tutti i cicli ausiliari volti a trattare situazioni di funzionamento

anomale, diagnostica, emergenza e riavvio a caldo o a freddo. In fase di progettazione concettuale molti comportamenti, o addirittura interi moduli o macchine, possono essere efficacemente modellati con blocchi funzionali software dotati di opportune interfacce input /output, realizzando modelli dettagliati solo per i moduli più critici.

Questa fase si conclude con la definizione di una combinazione di principi risolutivi che, integrando soluzioni meccaniche e software concettuali, realizzi i comportamenti necessari per il sistema adattivo di assemblaggio automatico.

Progettazione esecutiva

Nella fase di progettazione esecutiva si dimensionano o si progettano accuratamente gli apparati meccanici e si definiscono in dettaglio le relative strutture hardware e software di controllo. Si realizzano prototipi virtuali ad un sempre maggiore livello di dettaglio per la verifica e l'ottimizzazione delle combinazioni di soluzioni tecniche individuate nella precedente fase di progettazione concettuale. La sperimentazione virtuale dovrà essere accuratamente validata, [20], per arrivare ad uno stato di avanzamento del progetto che sia il più possibile privo di errori, dal momento che nella successiva fase di progettazione di dettaglio sarà più difficile introdurre modifiche. In questa fase sono molto importanti gli strumenti di analisi e visualizzazione grafica offerti dagli ambienti CAD 3D, per verificare i comportamenti progettati. In questa fase si devono simulare una grande quantità di diversi scenari virtuali, coinvolgendo tutte le diverse tecnologie. Nel caso del software, ad esempio, la stessa in fase di debug può essere vista come una sorta di modellazione, dal momento che il software reso eseguibile con la compilazione non sarà più comprensibile al progettista. La sequenza ed i risultati della simulazione (continua o step by step) sono normalmente visualizzati in un debugger, che mostra lo stato corrente di avanzamento nel codice appena sviluppato ed i risultati a monitor.

In certi casi si preferisce introdurre i comportamenti reali con un approccio di tipo Hardware in the Loop, [21, 22], connettendo alla simulazione uno o più controllori reali ed eventualmente anche i drive degli azionamenti. Questo approccio consente analisi ad un elevato livello di dettaglio del sistema, dal momento che il comportamento del componente reale utilizzato non è in alcun modo semplificato dal processo di prototipazione virtuale. Si tratta comunque di una tecnica che deve intervenire in una fase avanzata del processo di progettazione.

Progettazione di dettaglio

In questa fase finale si definiscono gli ultimi dettagli del sistema automatico adattivo fino al completamento delle distinte base di tutti gli elementi meccanici, elettrici ed elettronici ed alla redazione della documentazione tecnica del sistema. Sulla base delle simulazioni virtuali i progettisti meccanici devono completare l'ottimizzazione del sistema, gli ultimi dettagli del layout generale, la scelta della componentistica commerciale, dei materiali, delle proprietà superficiali e delle tolleranze. L'ingegneria del software in quest'ultima fase provvede al commissioning del software di controllo dell'intero sistema, progettando e verificando tutti i cicli ausiliari di diagnostica, di avviamento a caldo e a freddo e di arresto e le funzioni di comunicazione tra i diversi apparati collegati in rete.

Nella progettazione di dettaglio le diverse aree disciplinari devono lavorare con strumenti separati dal momento che questi sono stati perfezionati in decenni di sviluppo. In questa fase si deve quindi ricorrere ad una simulazione di tipo Hardware in the Loop. Mentre da un lato il software di controllo viene valutato nel suo reale comportamento, la

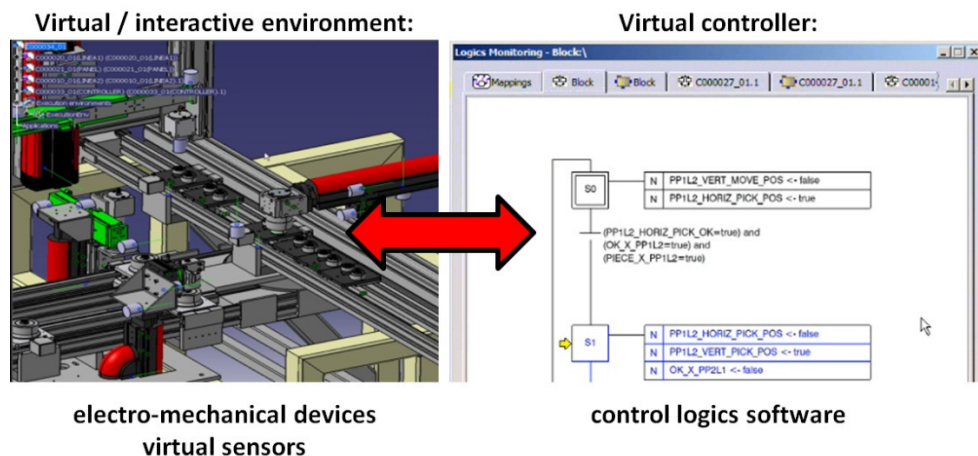


Figura 1. Simulazione in ambiente virtuale.

meccanica rimane definita, prima della messa in servizio e del collaudo finale, solo nell'ambiente virtuale di simulazione. Le simulazioni virtuali del comportamento meccanico possono però essere raffinate, se necessario, mediate l'introduzione di parametri di funzionamento più accurati, ottenuti anche tramite sperimentazione reale.

4. STRUMENTI DI PROGETTAZIONE INTEGRATA

L'integrazione di simulazioni meccaniche, elettriche ed elettroniche e delle logiche di controllo è stata ampiamente studiata nella progettazione di macchine, [21, 23], mentre un minor numero di applicazioni hanno come oggetto i sistemi automatici di assemblaggio. Nonostante la crescente diffusione della simulazione e programmazione off-line di sistemi robotizzati, [24], la programmazione e la validazione off-line dei PLC è ancora limitata a progetti pilota, tutti basati sulla tecnologia del Virtual Commissioning (VC), [12, 25]. In questo lavoro sono stati scelti gli strumenti del VC perchè efficientemente organizzati in ambienti interdisciplinari in cui le aree meccanica, elettro-meccanica e software possono comunicare e scambiare informazioni, creare moduli funzionali integrati e formalizzare basi di conoscenza interdisciplinari.

Vi sono due possibili approcci per la realizzazione di un prototipo di un intero sistema automatico con gli strumenti del VC. L'approccio più semplice consiste nel realizzare una simulazione completa in un ambiente virtuale, come mostrato in Fig. 1, in cui vengono integrati i comportamenti degli apparati meccanici ed elettronici e del software di controllo. La simulazione del comportamento integrato di un sistema può così essere realizzata su un comune PC e consente di adottare tempi di simulazione perfettamente sincronizzati tra modelli anche molto complessi. Gli ambienti di simulazione sono molto aperti e permettono un elevato grado di flessibilità nella definizione di modelli a diversi livelli di dettaglio, per quanto riguarda la meccanica degli azionamenti, gli schemi elettrici, i circuiti pneumatici ed idraulici ed i sensori. D'altra parte questo si riflette in lunghi tempi di sviluppo, per cui risulta fondamentale la creazione di database di librerie di modelli parametrici, da riutilizzare velocemente ed integrare in nuove simulazioni, così come normalmente realizzato dall'ingegneria del software. Il limite principale di questo approccio è la necessità di sviluppare il software di controllo in ambienti con funzionalità semplificate

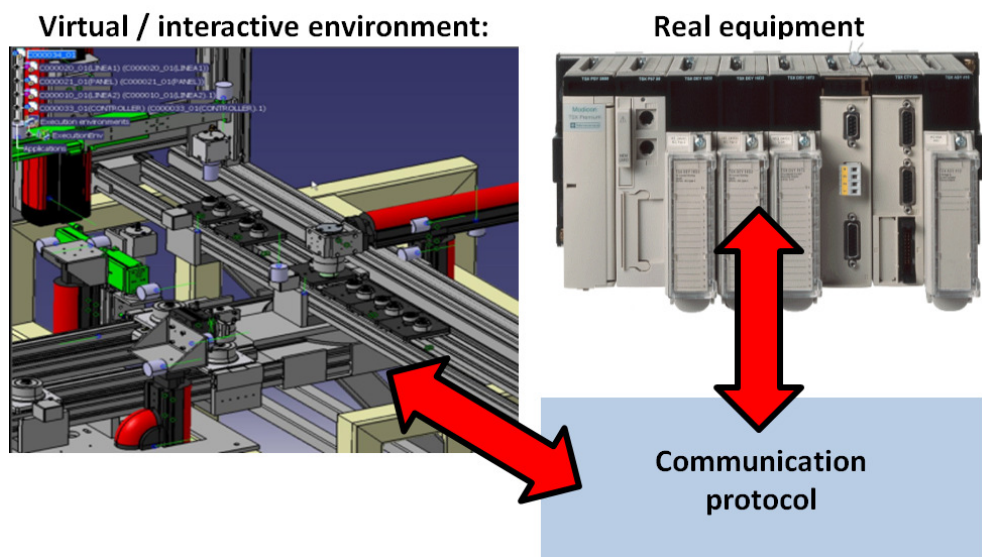


Figura 2. Simulazione Hardware in the Loop.

rispetto agli strumenti di sviluppo proprietari dei costruttori di controllori industriali. La difficoltà di traduzione verso gli ambienti proprietari da cui eseguire infine il download dell'applicazione su un controllore fa sì che sia conveniente sviluppare solamente una parte del software reale, e tipicamente un comportamento privo dei cicli ausiliari di avviamento e freddo, di emergenza e di restart e di diagnostica.

Il secondo approccio consiste nel realizzare una simulazione Hardware in the Loop, [21], come mostrato in Fig. 2, in cui solo una parte del sistema viene simulato, mentre una o più parti vengono eseguite tramite l'hardware che equipaggerà il sistema reale. In generale si tratta delle apparecchiature hardware e software reali di controllo, and esempio PLC. Questo approccio consente di utilizzare gli ambienti di sviluppo proprietari delle diverse tecnologie, conservando per intero la base di conoscenza costituita dalle librerie software, nonché tutti gli strumenti di configurazione hardware e di diagnostica. Se da un lato con questo approccio si ha il grande vantaggio di utilizzare il software così come verrà realmente eseguito, dall'altro la simulazione integrata risente delle prestazioni del sistema di comunicazione tra controllore reale ed ambiente di simulazione. Questi infatti possono non essere perfettamente sincronizzati a causa della comunicazione tramite un server OPC (OLE – Object Linking and Embedding – for Process Control), che regola le comunicazioni dei segnali dell'ambiente di simulazione con i dispositivi di controllo di diversi produttori, inclusi i sistemi HMI (Human Machine Interface).

Questi strumenti consentono quindi di colmare la separazione tra le diverse aree dell'ingegneria con la ricerca, la valutazione, ed infine la verifica dei risultati di sinergie interdisciplinari, a patto che vengano opportunamente guidati da un nuovo metodo di progettazione.

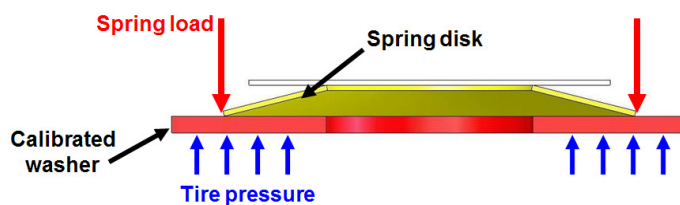


Figura 3. Principio di funzionamento del sensore di pressione.

5. PROGETTAZIONE INTEGRATA DI UN SISTEMA ADATTIVO DI ASSEMBLAGGIO DI SENSORI AUTOMOTIVE

La prototipazione virtuale è stata adottata come strumento di sviluppo di alcune soluzioni tecniche adattive progettate per produrre ed assemblare sensori di pressione automotive, [26]. Il principio di funzionamento del sensore, come mostrato in Fig. 3, è basato su una rosetta calibrata su cui agisce una molla a tazza: l'area della corona circolare della rosetta è soggetta alla pressione differenziale tra l'aria all'interno dello pneumatico e quella esterna e la forza risultante deve essere controbilanciata dalla molla a tazza. La molla è progettata per avere una caratteristica fortemente nonlineare, con bassi valori di costante elastica in prossimità del carico di attivazione desiderato per il sensore. Alla pressione di attivazione, rilevata con un'ottima precisione, il sensore attiva un segnale visivo che indica lo stato di pressione effettiva dello pneumatico. Il modulo riprogettato in questo case study è responsabile dell'assemblaggio della molla a tazza e della rosetta calibrata.

La variabilità dei parametri produttivi ritenuti più critici è stata accuratamente analizzata e modellata sin dalle prime fasi di definizione delle specifiche e progettazione concettuale del sistema di assemblaggio. La modellazione degli effetti di tali variabilità sulle prestazioni finali del sensore di pressione ha evidenziato alcune criticità. In particolare è necessario variare frequentemente il diametro della rosetta calibrata, a causa della richiesta di frequenti cambi di lotto di produzione (ogni lotto ha un valore diverso della soglia di pressione, e quindi un diametro diverso della rosetta), richiedendo lo svuotamento manuale ed il nuovo riempimento degli alimentatori vibranti con tempi di set up molto lunghi. L'analisi FEM sulla molla a tazza ha inoltre mostrato come la curva elastica nonlineare cambi con il cubo dello spessore (dell'ordine di 0.1mm). Poiché il processo di tranciatura delle molle è difficile da controllare a costi accettabili e comporta in definitiva grandi variazioni della curva elastica della molla, è stato necessario intervenire a valle, progettando un sistema adattivo rispetto alla variazione della caratteristica elastica della molla.

I comportamenti del sistema sono stati via via modellati tramite prototipi virtuali con il software Delmia Automation della Dassault Systèmes. In questo ambiente virtuale si sviluppano i modelli meccanici CAD 3D, i relativi vincoli cinematici e le logiche di comportamento. Una volta integrati con tali logiche i prototipi virtuali diventano smart devices, ovvero moduli intelligenti in grado di comunicare tra di loro e con il controllore tramite porte di IOs, come mostrato in Fig. 4. Questi prototipi virtuali dei sistemi elettromeccanici e pneumatici vengono inoltre integrati grazie a specifici strumenti di modellazione con attuatori programmabili, altre comunicazioni tramite IOs, sensori digitali e analogici, HMI, controllori e logiche di controllo.

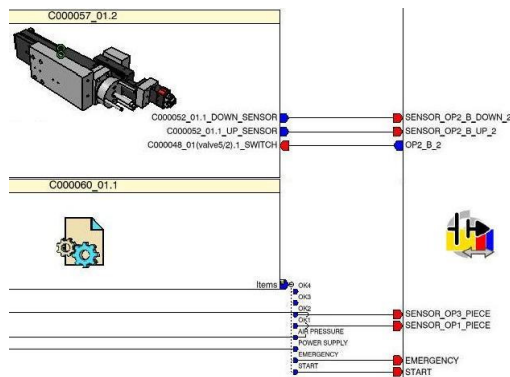


Figura 4. Comunicazione di uno smart device tramite porte IOs.

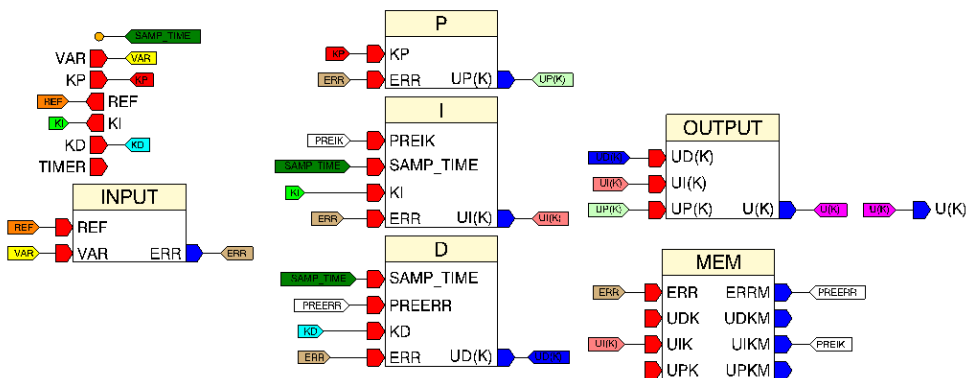


Figura 5. Modello FBD di un regolatore PID.

Nel presente lavoro l'utilizzo degli strumenti di prototipazione virtuale è stato anticipato alle prime fasi di sviluppo del sistema. L'ambiente è stato utilizzato in parallelo da ingegneri meccanici e da progettisti software e la stessa logica comportamentale degli apparati meccanici è stata definita tramite i linguaggi dello standard IEC 61131-3, per una migliore integrazione tra le diverse aree di conoscenza. In particolare i comportamenti dinamici dei sistemi meccanici e degli attuatori programmabili sono stati modellati con il linguaggio FBD (Function Block Diagram), mentre per i comportamenti logici, tipici di una descrizione a stati discreti, è stato utilizzato il linguaggio SFC+ (Sequential Function Chart).

In Fig. 5 si mostra un esempio di modellazione tramite FBD del controllo PID di un attuttore programmabile, utilizzato nella valutazione concettuale di soluzioni attuate elettricamente. Un FBD è un diagramma composto da blocchi funzionali che descrivono relazioni logiche, algebriche o ausiliarie (ad esempio cambio di formato) tra variabili di input e di output (IOs). Le porte IOs di blocchi diversi sono connesse da linee orientate da sinistra verso destra, sostituibili da etichette per chiarezza di rappresentazione. Due porte connesse tra di loro devono essere dello stesso formato. Una porta di input può essere collegata ad un solo output, mentre una porta di output può essere collegata a più input.

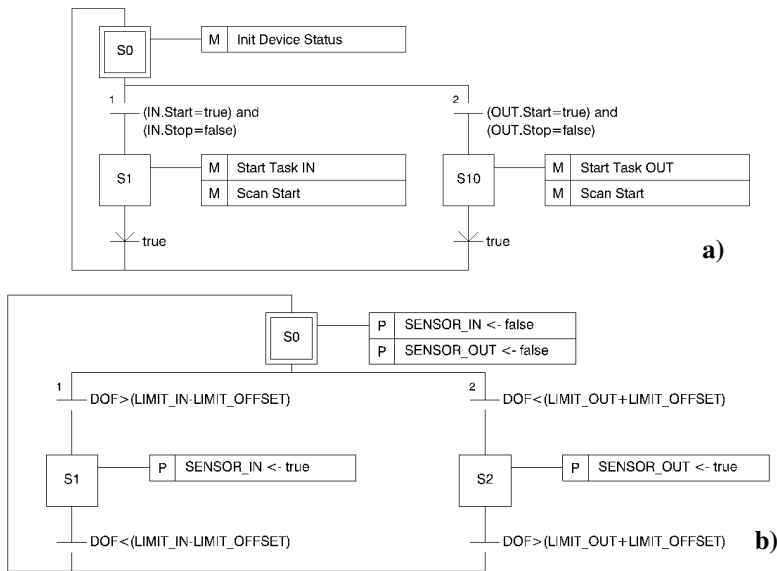


Figura 6. Modello SFC+ di un attuatore pneumatico sensorizzato.

Il modello del regolatore PID comprende blocchi per il processamento delle variabili IOs, i blocchi proporzionale, integrale e derivativo ed una funzione di memorizzazione dello stato e degli stati precedenti. La dinamica dell'attuatore controllato è stata anch'essa modellata tramite un FBD. Il modello parametrico è completamente configurabile con i parametri cinematici (corsa, posizioni iniziale e finale), dinamici (inerzie), elettrici (parametri del motore) e di controllo (trasduttore di posizione). Nello stesso FBD la dinamica del sistema è governata da blocchi funzionali algebrici, mentre altre condizioni, come le condizioni iniziali e l'arresto ai fine-corsa viene gestito da blocchi logici di confronto. Le funzioni inserite nei blocchi sono state scritte completamente oppure importate da funzioni definite in Matlab con un linguaggio simbolico.

In Fig. 6 si mostra un esempio di modellazione tramite SFC+ di un cilindro pneumatico di un'unità *pick and place*, soluzione poi adottata come modulo standard in molte soluzioni del sistema di assemblaggio. L'SFC+ è un linguaggio di programmazione grafica concettualmente simile al GRAFCET e alle reti di Petri e risulta molto efficiente nella modellazione di comportamenti descrivibili con sequenze (in serie, in parallelo o alternative) di stati discreti. Ad ogni stato possono essere associate delle azioni da eseguire, a loro volta definite con assegnazioni, equazioni e condizioni logiche di confronto. Gli stati sono connessi tra di loro mediante transizioni, attivabili con la verifica di opportune condizioni.

Anche il modello comportamentale SFC+ dell'attuatore pneumatico, Fig. 6 (a), è completamente parametrico ma i movimenti sono vincolati ad eseguire profili di velocità trapezoidali o triangolari. Questo modello virtuale comprende anche la modellazione dei sensori di prossimità per i finecorsa, Fig. 6 (b), per cui è possibile configurare il tipo, la posizione e la distanza di lettura. Questo tipo di modellazione è molto aperta e consente un gran numero di prove virtuali per la validazione del comportamento complessivo del

sistema, includendo la replica di un gran numero di scenari ipotetici di guasto, a cui il sistema deve reagire opportunamente.

I modelli parametrici così costruiti sono disponibili in file XML o nativi Delmia e quindi facilmente riutilizzabili per successive attività di progettazione o istanziabili più volte all'interno della stessa simulazione, anche con parametri diversi.

Le diverse stazioni operative, di volta in volta definite a livello concettuale e disposte nel layout con una certa approssimazione, sono state inizialmente modellate semplicemente con blocchi funzionali software e con schematizzazioni CAD 3D per definirne gli ingombri nello spazio. I modelli sono stati quindi man mano definiti ad un livello maggiore di dettaglio, a partire da quelli più critici per conseguire le specifiche del sistema. Per ogni smart device simulato in ambiente virtuale sono state definite le necessarie interfacce IOs per il trasferimento di materia, energia ed informazioni sia a livello hardware (con i moduli confinanti e con le HMI) che software (con il controllore virtuale). Per comprimere i tempi di sviluppo sono stati riutilizzati ampiamente i modelli precedentemente sviluppati e disponibili in libreria, con le rispettive informazioni sulle reali prestazioni ottenibili. Anche i blocchi funzionali PLC delle logiche di controllo sono stati riutilizzati e modificati per valutare rapidamente la fattibilità di ogni soluzione o variante di progetto e per verificare i comportamenti integrati di meccanica e software.

I modelli concettuali delle diverse soluzioni tecniche sono stati assemblati in un modello di sistema complessivo, via via integrato con i comportamenti adattivi ad un sempre maggiore livello di dettaglio. Il sistema adattivo è stato quindi sviluppato progressivamente con un approccio Top-Down tramite l'introduzione di tutti i necessari dispositivi meccanici, i sensori ed i sistemi mecatronici programmabili, con le necessarie capacità di acquisizione, processamento e comunicazione di informazioni.

L'architettura del sistema adattivo è stata definita secondo il criterio di separazione degli apparati responsabili del comportamento adattivo dal sistema principale. Sono stati quindi progettati due diversi anelli di trasporto. L'utilizzo di due anelli con velocità di avanzamento differenti consente di ottimizzare il sistema principale senza risentire delle problematiche e delle eventuali inefficienze legate ai comportamenti adattivi. I due anelli di trasporto sono poi stati interfacciati con alcune stazioni pick and place, come mostrato in Fig. 7.

Visti i punti deboli del sistema precedentemente evidenziati, si è deciso di testare in linea ogni singola molla nell'anello adattivo, per poter ottimizzare l'area superficiale della rosetta accoppiata in funzione della curva elastica registrata. Per questo comportamento adattivo sono state definite, come in [5], le attività di *environment sensing*, *goal modification* e *behaviour adaptation*. È stata introdotta una stazione di test in cui si acquisisce la curva elastica per ogni singola molla (environment sensing). Le molle a tazza sono testate una ad una per cinque volte, per formare, compatibilmente con i tempi ciclo richiesti, una minima base statistica per la successiva elaborazione. Quindi la logica di controllo determina una curva di carico media in funzione della compressione e di conseguenza calcola il diametro (goal modification) con cui tagliare la rosetta (dal diametro dipende l'area di azione della pressione) per ottenere la specifica soglia di attivazione prevista per il lotto in lavorazione. Infine le rosette grezze sono alimentate da un alimentatore a tazza vibrante in una stazione laser in cui vengono singolarmente tagliate a misura (behaviour adaptation).

Un opportuno sistema di tracking dei pallet e di memorizzazione delle operazioni assicura che per ogni sensore vengano assemblati correttamente la molla e la rosetta appositamente adattata. Mentre la posizione dei singoli sottoassiemi (sensori parzialmente

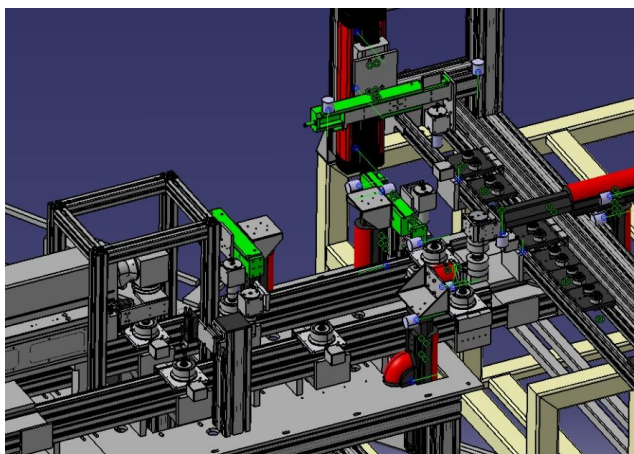


Figura 7. Sincronizzazione dei diversi anelli di trasporto tramite unità pick and place.

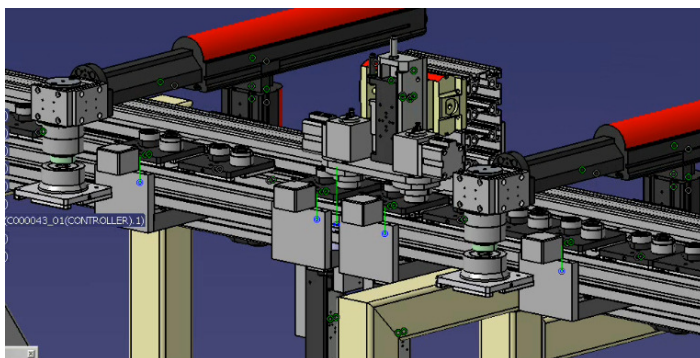


Figura 8. Sistema di trasporto asincrono a pallet liberi.

assemblati) è seguita lungo le varie stazioni operative, il laser taglia precisamente a misura una rosetta sulla base della soglia di pressione prevista per il lotto in lavorazione e della curva elastica misurata per la specifica molla a tazza. Per il controllo CN del laser è inoltre previsto un ciclo opportuno per recuperare i giochi ed ottenere una maggiore ripetibilità di taglio.

La prototipazione virtuale dei sistemi di trasferimento asincroni a pallet liberi è stata particolarmente accurata sia dal punto di vista meccanico che software. Con questo sistema, mostrato in Fig. 8, i pallet vengono fatti continuamente avanzare fino a che non vengono ostacolati da un pallet precedente o dai fermi controllati dal PLC. In corrispondenza di ogni stazione vi sono due fermi in successione ed un sistema di bloccaggio e riferimento del pallet. Mentre la stazione operatrice esegue le operazioni sul pallet bloccato precisamente, il fermo a monte impedisce che i pallet in arrivo collidano col pallet in lavorazione, generando spostamenti o vibrazioni impreviste. Per la modellazione virtuale di questi comportamenti nelle diverse condizioni di lavoro sono stati accoppiati blocchi FBD ed SFC+. La prototipazione di questi sistemi ha richiesto un'elevata modularizzazione per consentire di inserire i diversi pallet come elementi indipendenti ma interagenti tra di loro

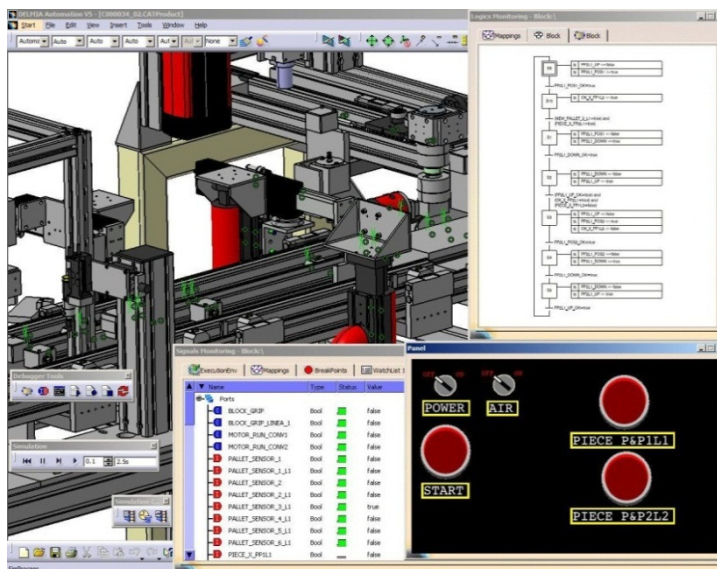


Figura 9. Simulazione virtuale mediante Delmia Automation.

da un punto di vista puramente fisico. La prototipazione virtuale del sistema di trasporto ha consentito di ottimizzare lo spazio libero tra diverse stazioni, in cui si accumulano buffer di pallet in attesa, ed il numero stesso dei pallet in circolazione. In genere infatti l'efficienza del sistema può diminuire sia per un numero di pallet troppo scarso che troppo elevato. Un numero ottimale di pallet al contrario consente di saturare i colli di bottiglia del sistema.

Il comportamento del sistema adattivo è stato infine completamente implementato in un prototipo virtuale integrato con tutti gli smart devices e logiche di controllo. Per eseguire il debug del software e per validare e verificare le prestazioni dei sistemi implementati sono state utilizzate sia simulazioni continue nel tempo che step-by-step (Fig. 9). In questo modo è stato possibile valutare i movimenti, le collisioni, il flusso delle operazioni, il funzionamento della strategia di sensorizzazione ed il valore dei segnali nel tempo. I comportamenti sono stati rapidamente ottimizzati grazie all'interazione dei progettisti con l'ambiente di simulazione virtuale.

6. CONCLUSIONI

In questo lavoro sono state studiate le diverse fasi di progettazione dei sistemi adattivi di assemblaggio automatico per individuare metodi e tecniche finalizzati a superare alcuni limiti della tradizionale progettazione sistematica e ad integrare le diverse aree disciplinari. L'approccio descritto è basato sull'utilizzo della prototipazione virtuale come vero e proprio strumento di progettazione. In tutte le fasi di sviluppo sistematico i comportamenti adattivi vengono progettati integrando le diverse aree tecnologiche coinvolte. I prototipi virtuali comprendono quindi i modelli CAD 3D, i comportamenti meccanici, elettrici e pneumatici, i processi fisici, le diverse tipologie di sensori, l'hardware ed il software di controllo, le interfacce uomo macchina e qualunque altro tipo di modellazione funzionale alla simulazione. I modelli si scambiano flussi di materia, energia ed informazioni mediante porte IOs di diversi formati, supportando così efficaci strategie di modularizzazione.

Il metodo, le tecniche e gli strumenti descritti sono stati valutati su diversi casi studio, tra i quali è stata descritta la progettazione di un sistema automatico per la produzione e l'assemblaggio di sensori automotive. In questo caso studio, la definizione iniziale delle specifiche prestazionali del prodotto da assemblare è stata integrata con informazioni sulla variabilità dei parametri di lavoro. Alcuni prototipi concettuali sono stati utilizzati per valutazioni preliminari e per individuare i punti deboli del sistema da superare con soluzioni adattive. Grazie alla sperimentazione virtuale dell'evoluzione dei comportamenti nel tempo è stato possibile definire le logiche comportamentali e le strategie di sensorizzazione ed attuazione necessarie per la loro implementazione. L'attività ha portato ad un notevole incremento delle prestazioni del sistema in termini di qualità ottenuta, capacità produttiva (grazie soprattutto all'eliminazione dei fermi macchina), autonomia, riduzione dell'intervento umano e flessibilità produttiva. Infine la simulazione complessiva del sistema adattivo ha mostrato come sia possibile sviluppare in tempi molto rapidi un prototipo virtuale molto complesso, grazie anche alla creazione di librerie di modelli virtuali integrati. Grazie al metodo di progettazione basato sulla prototipazione virtuale è stato possibile migliorare la qualità del software ed al tempo stesso ridurre notevolmente i tempi di messa a punto. Questi risultati, in linea con altre recenti ricerche, sono ottenuti grazie alla possibilità di eseguire test intensivi, anche nelle più critici scenari operativi, che spesso non potrebbero essere ricreati nella realtà senza il rischio di danneggiare i macchinari.

L'attività di ricerca futura prevede l'approfondimento dei modelli matematici con cui sono creati i moduli parametrici dei sistemi e la connessione con altri software di modellazione ed ottimizzazione di sistemi automatici. Si dovranno poi approfondire le metodologie dell'ingegneria meccanica nella definizione ai diversi livelli di dettaglio dei comportamenti adattivi, a partire dalla definizione delle specifiche con nuovi linguaggi di formalizzazione capaci di descrivere l'incertezza delle condizioni operative, fino alla progettazione vera e propria con un approccio sintetico e sistemico.

RINGRAZIAMENTI

Questo progetto è stato realizzato grazie al supporto delle strutture operative del laboratorio regionale dell'Emilia Romagna per l'Alta Meccanica INTERMECH-SIMECH, oggi Centro Interdipartimentale per la Meccanica Avanzata e la Motoristica INTERMECH-Mo.Re. dell'Università di Modena e Reggio Emilia.

BIBLIOGRAFIA

- [1] De Toni, A., and Tonchia, S., 1998. "Manufacturing flexibility: a literature review", *Int. J. Prod. Res.*, **36**(6), pp. 1587-1617.
- [2] ElMaraghy H.A., 2006. "Flexible and reconfigurable manufacturing systems paradigms", *Int J Flex Manuf Syst*, **17**, pp. 261-276.
- [3] Beach, R., Muhlemann, A.P., Price, D.H.R., Paterson, A., and Sharp, J.A., 2000. "A review of manufacturing Flexibility", *European Journal of Operational Research*, **122**, pp. 41-57.
- [4] Gupta, Y.P., and Goyal, S., 1989. "Flexibility of manufacturing systems: concepts and measurements", *European Journal of Operational Research*, **43**, pp. 119-135.
- [5] Burmester, S., Giese, H., Münch, E., Oberschelp, O., Klein, F., and Scheideler, P., 2008. "Tool support for the design of self-optimizing mechatronic multi-agent systems", *Int. J. Softw. Tools Technol. Transf.*, **10**, pp. 207-222.

- [6] Li, P.Y., and Horowitz, R., 1997. "Self-optimizing control". Proceedings of the 36th Conference on Decision and Control, SanDiego, California, USA.
- [7] Cao, Y., ElMaraghy, H.A., and Azab, A., 2007. "Reconfigurable control structure for robots in assembly", *J. Intell. Robot. Syst.*, **50**, pp. 419-439.
- [8] Cheng, B.H.C., de Lemos, R., Giese, H., Inverardi, P., and Magee, J., 2009. *Software Engineering for Self-Adaptive Systems: A Research Roadmap*, Springer-Verlag, Berlin, pp. 1-26.
- [9] Isermann, R., 1981. "Advanced methods of process computer control for industrial processes", *Computers in Industry*, **2**, pp. 59-72.
- [10] Pahl, G., and Beitz, W., 1996. *Engineering Design - A Systematic Approach*, Springer-Verlag, Berlin.
- [11] Cassandras, C.G. and Lafortune, S., 2008. *Introduction to Discrete Event Systems*, Springer Science + Business Media, New York, NY, USA.
- [12] Reinhart, G., and Wunsch, G., 2007. "Economic application of virtual commissioning to mechatronic production systems", *Prod. Eng. Res. Devel.*, **1**, pp. 371-379.
- [13] Andrisano, A.O., Leali, F., Pellicciari, M., and Vergnano, A., 2009. "Engineering Method for Adaptive Manufacturing System Design", *International Journal on Interactive Design and Manufacturing*, **3**(2).
- [14] AberdeenGroup Inc., 2006. *The Mechatronics System Design Benchmark Report*, AberdeenGroup White Paper, Boston, MA, USA.
- [15] Hewit, J., 1996. "Mechatronics design: the key to performance enhancement", *Robot. Auton. Syst.*, **19**(2), pp. 135-142.
- [16] Qureshi, N.A., and Perini, A., 2009. "Engineering Adaptive Requirements", Proceedings of ICSE 2009 Workshop on Software Engineering for Adaptive and Self-Managing Systems (SEAMS 2009), Vancouver, BC, Canada.
- [17] Berkovich, M., Leimeister, J.M., and Krcmar, H., 2009. "An empirical exploration of requirements engineering for hybrid products", Proceedings of 17th European Conference on Information Systems, Verona, Italy.
- [18] Andersson, J., de Lemos, R., Malek, S., and Weyns, D., 2009. "Modeling dimensions of self-adaptive software systems", in: Cheng, B.H.C., de Lemos, R., Giese, H., Inverardi, P., Magee, J., *Software Engineering for Self-Adaptive Systems: A Research Roadmap*, Springer, Heidelberg, pp.27-47.
- [19] Pelz, G., 2003. *Mechatronic Systems -- Modelling and Simulation with HDLs*, John Wiley & Sons Ltd, Chichester, UK.
- [20] Kleijnen, J.P.C., 1995. "Verification and validation of simulation models", *European Journal of Operational Research*, **82**(2), pp. 145-162.
- [21] Kanai, S., Miyashita, T., and Tada, T., 2007. "A multi-disciplinary distributed simulation environment for mechatronic system design enabling hardware-in-the-loop simulation based on HLA", *Int. J. Interact. Des. Manuf.*, **1**, pp. 175-179.
- [22] Andrisano, A.O., Leali, F., Pellicciari, M., and Vergnano, A., 2008. "Mechatronic Design of Adaptive Manufacturing Systems", Proceedings of the International Conference IDMME - Virtual Concept 2008, Beijing, China.
- [23] Broenink, J.F., and Kleijn, C., 1999. "Computer-aided design of mechatronic systems using 20-SIM 3.0", Proceedings of WESIC'99, 2nd Workshop on European Scientific and Industrial Collaboration, Newport, UK.

- [24] Andrisano, A.O., Leali, F., Pellicciari, M., Pini, F., and Vergnano, A., 2007. "Design methods for intelligent robotic deburring cells", Proceedings of the 6th International Conference on Processing and Manufacturing of Materials, Salerno, Italy.
- [25] ARC Advisory Group, 2006. *V5 Automation Platform: Merging Digital Manufacturing with Automation*, ARC White Paper, Dedham, MA, USA.
- [26] Barbanti, G., 2004. "System for predetermining critical value of device for inspecting radial deformation of tire, composed of member for controlling pulling force caused by tire's pressure, belleville washer, and member for controlling state of equilibrium against pulling force. Patent IPC: B60C23/06.

LA MECCATRONICA PER L'ASSEMBLAGGIO FLESSIBILE

Aldo Rossi

*Dipartimento di Innovazione Meccanica e Gestionale (DIMEG),
Università degli Studi di Padova
E-mail: aldo.rossi@unipd.it*

Giulio Rosati

*Dipartimento di Innovazione Meccanica e Gestionale (DIMEG),
Università degli Studi di Padova
E-mail: giulio.rosati@unipd.it*

Stefano Cenci

*Dipartimento di Innovazione Meccanica e Gestionale (DIMEG),
Università degli Studi di Padova
E-mail: stefano.cenci@unipd.it*

Andrea Carli

*Dipartimento di Innovazione Meccanica e Gestionale (DIMEG),
Università degli Studi di Padova
E-mail: andrea.carli@mechatronics.it*

Luca Zanotti

*Sierra SpA
Isola Della Scala (VR)*

Massimo Mantovani

*Sierra SpA
Isola Della Scala (VR)*

Alberto Foroni

*Sierra SpA
Isola Della Scala (VR)*

Valerio Giordano Riello

*Sierra SpA
Isola Della Scala (VR)*

Sommario. Viene illustrata la metodologia di progetto di una cella robotizzata flessibile per assemblaggio automatico dei componenti “curva” sulla sommità dei tubi costituenti un circuito idraulico. Partendo dal ciclo produttivo preesistente e conoscendone le specifiche, è affrontato il progetto dell’impianto automatico, definendone preliminarmente il layout e individuando i sottosistemi che compongono l’architettura della cella robotizzata. L’impianto automatico è dotato di magazzino utensili interfacciato con un sistema di bowl-feeders per l’orientazione e l’alimentazione delle curve da montare. Gli end-effector per il robot S.C.A.R.A., depositati a magazzino, sono provvisti ciascuno di un caricatore, il quale viene riempito con i componenti da montare dal sistema di alimentazione. Una tavola rotante elettrica, dotata di maschere di fissaggio, costituisce il sistema di trasporto interno alla cella e permette il carico/scarico e la staffatura del prodotto durante la fase di montaggio. Il robot è dotato di un cambio utensili a comando pneumatico, che permette al manipolatore di dotarsi, di volta in volta, di uno degli end-effector presenti a magazzino, a seconda delle necessità richieste dal ciclo di assemblaggio. I prodotti assemblabili dalla cella robotizzata possono variare in dimensioni, numero e tipologia di curve da montare, perciò è presente un sistema di visione in grado di riconoscere la tipologia di prodotto presente sulla tavola rotante, individuarne eventuali difetti di fabbricazione, definire i centri tubo sui quali montare le curve ed effettuare una foto dell’assemblaggio eseguito, verificando la qualità e la completezza dell’operazione, permettendo inoltre la tracciabilità dei prodotti assemblati. Il software di gestione dell’impianto automatico effettua il controllo e il monitoraggio dei sottosistemi costituenti la cella, consentendo l’assemblaggio di prodotti differenti senza che si renda necessaria alcuna riconfigurazione o riprogrammazione della cella di lavoro, e si occupa di definire la sequenza ottimale di assemblaggio delle curve, minimizzando il tempo dell’operazione. Inoltre, il software è stato progettato per interfacciarsi con il sistema gestionale dello stabilimento produttivo.

Parole Chiave: Assemblaggio Automatico, Cella Robotizzata Flessibile, End-Effector, Sistema di Visione.

1. INTRODUZIONE

L’attuale tendenza del mercato di richiedere prodotti personalizzabili e in lotti limitati, per via del loro ridotto ciclo di vita, pone numerosi problemi nell’automatizzare le fasi di un ciclo produttivo. Bisogna senz’altro sottolineare che, mentre le tecnologie di produzione tradizionali oggi sono dotate di un elevato livello di automazione, questa, seppur presente, rimane inferiore nella fase di assemblaggio. Si può inoltre evidenziare che la razionalizzazione dell’assemblaggio ha attualmente acquistato una grande importanza, se non altro perché la sua quota relativa ai tempi e ai costi di produzione rappresenta una parte ragguardevole con tendenze sempre più rilevanti [1]. I sistemi di assemblaggio automatico devono possedere un livello di flessibilità correlato ai quantitativi di produzione richiesti, all’ampiezza dei lotti, al numero dei modelli di prodotto, delle loro varianti e della loro complessità [2]. Lo sviluppo di questi sistemi è stato posteriore a quello verificatosi nel campo della lavorazione di macchina anche in relazione alla maggiore complessità del problema [3, 4, 5, 6, 7, 8].

Il ciclo produttivo tradizionale dei prodotti prevede diverse fasi di assemblaggio tutte eseguite da operatori manuali. L’assemblaggio così realizzato si basa su risorse assolutamente flessibili, quali sono gli operai assemblatori, ma caratterizzate da limitata

ripetibilità e discrezionalità nella valutazione di parametri qualitativi. L'idea di introdurre un impianto automatico di montaggio, oltre a garantire un determinato tempo ciclo, la standardizzazione del livello qualitativo [9, 10], la certificazione e la tracciabilità del prodotto realizzato, consente d'altra parte di prelevare personale da quelle attività ripetitive e noiose per destinarlo a mansioni maggiormente valorizzanti. Questo aspetto inoltre concorda con l'attuale esigenza di ridurre l'incidenza del costo della manodopera sul costo finale del prodotto.

Il progetto e la successiva realizzazione della cella robotizzata sono il risultato di un lavoro che è partito dall'analisi del prodotto e del suo ciclo di fabbricazione, potendo in questo modo definire le specifiche necessarie in termini di sottosistemi presenti, disposizione planimetrica, performance, integrazione con il ciclo fabbrica e grado di flessibilità. Dopo avere valutato la fattibilità economica del progetto e la sua convenienza si è passati alla progettazione dei sottosistemi che costituiscono l'impianto automatico. In particolare si è studiata la modalità con cui alimentare i componenti da assemblare, si è definito il sistema di carico/scarico e staffatura dei prodotti durante la fase di montaggio, si è operata la scelta della tipologia e delle caratteristiche del robot industriale presente nell'impianto. Parallelamente ci si è dedicati al progetto degli end-effector e del loro magazzino di deposito, oltre alla definizione del sistema di visione, con la scelta dei suoi componenti e delle loro caratteristiche [11, 12]. Un aspetto rilevante è stato quello relativo alla programmazione del software di gestione della cella, il quale si interfaccia con il sistema di visione, con il controllore del robot e con il gestionale dello stabilimento produttivo.

La fase di progetto è stata seguita da una fase di prototipazione della cella, durante la quale ci si è avvalsi di macchina FDM per il rapid prototyping e quindi la veloce realizzazione dell'end-effector e del sistema di alimentazione. Avendo realizzato il prototipo è stato possibile eseguire i test per verificare e ottimizzare le performance del sistema. Infine l'impianto automatico è stato realizzato e collaudato presso lo stabilimento di un'azienda leader nel settore.

2. LA CELLA ROBOTIZZATA DI ASSEMBLAGGIO

La cella di lavoro robotizzata ha il compito di assemblare in maniera autonoma, intelligente e automaticamente riconfigurabile, gli elementi curva sui tubi costituenti il circuito idraulico dei prodotti, chiudendo così il percorso previsto per il fluido primario. La cella robotizzata è rappresentabile come un sistema che ha come input il prodotto da assemblare e i relativi disegni CAD, contenenti le specifiche di assemblaggio, e come output il prodotto assemblato con una serie di dati relativi all'operazione di assemblaggio avvenuta. Tra questi dati vi sono delle informazioni statistiche relative al tipo e numero di prodotti assemblati e non assemblati, a causa di difetti rilevati dal sistema di visione, quali e quanti di questi difetti sono stati rilevati e la gestione degli ordini.

Specifiche dell'impianto automatico di assemblaggio delle curve

L'operazione del ciclo produttivo che si vuole automatizzare è quella di inserimento delle curve con anelli nei tubi di un circuito idraulico. Alla luce delle attuali tendenze del mercato, che richiede la produzione di lotti di prodotti anche di bassa numerosità, dovendo quindi l'impianto avere la massima di versatilità e rapidità di riconfigurazione, si è deciso di

progettare una cella robotizzata di assemblaggio. Infatti il mix produttivo è tale da non giustificare la costruzione di una macchina automatica speciale. Anche effettuando considerazioni sulle varianti dei prodotti da assemblare e sulle caratteristiche di flessibilità richieste all'impianto, nonché valutazioni relative alla complessità del prodotto ed alle dimensioni dei lotti, si è indirizzati verso la progettazione di una cella robotizzata.

In una cella robotizzata destinata ad operare questo tipo di assemblaggio è essenziale avere un sistema di visione industriale che permetta di conoscere la posizione dei tubi nei quali deve entrare la curva. Ulteriori specifiche fondamentali di progetto individuate per l'impianto automatico sono tempo di inserimento della singola curva e l'accuratezza del posizionamento, stabilita in fase di progetto preliminare in ± 0.1 mm. La cella allo studio, inoltre, per poter operare in maniera efficiente all'interno della linea di produzione, deve soddisfare alcuni requisiti prestazionali e alcune necessità di produzione:

- **Interfacciamento:** la cella di lavoro deve utilizzare come elemento in ingresso il disegno CAD del prodotto da assemblare già presente nei sistemi informativi aziendali, provvedendo in maniera automatizzata all'acquisizione delle caratteristiche geometriche del prodotto ed alla generazione del relativo ciclo di assemblaggio, senza richiedere intervento alcuno da parte dell'operatore.
- **Velocità:** l'operazione eseguita all'interno della cella deve risultare più veloce della corrispondente lavorazione manuale. Tale requisito implica un'acquisizione ed un'elaborazione delle immagini ottimizzate e una pianificazione delle traiettorie di inserimento per il robot specifica per ogni prodotto.
- **Rilevazione difetti:** alla cella si richiede non solo di inserire correttamente e velocemente le curve nei tubi, ma al tempo stesso di individuare e quantificare eventuali difetti presenti alle imboccature dei tubi. Questo requisito ha almeno tre funzionalità. La prima è legata all'inserimento delle curve: in presenza di un difetto grave all'imboccatura di un tubo, l'identificazione dell'anomalia permette di evitare danneggiamenti all'organo terminale del robot ed al prodotto. Questo tipo di riconoscimento consente inoltre un'eventuale riparazione nella zona difettosa prima delle lavorazioni successive. Infine, le informazioni ricavate possono essere utili per un'analisi statistica sulla distribuzione dei difetti.

Architettura della cella robotizzata

Lo scopo della pianificazione del layout consiste nel disporre opportunamente la cella di assemblaggio, per quanto riguarda il flusso dei prodotti e la messa a disposizione delle curve da assemblare, in funzione dello spazio presente e rispettando le esigenze del personale in servizio. L'intera disposizione deve essere progettata in modo che il flusso dei semilavorati nella cella possa essere integrato nello svolgimento complessivo della fabbricazione. Devono inoltre essere considerati la modalità di messa a disposizione delle curve da assemblare ed i relativi percorsi. In particolare, la messa a disposizione mediante sistemi automatici voluminosi, se possibile, dovrebbe essere effettuata da un solo lato. Infine devono essere soddisfatte tutte le prescrizioni indicate nelle direttive di sicurezza per i luoghi di lavoro.

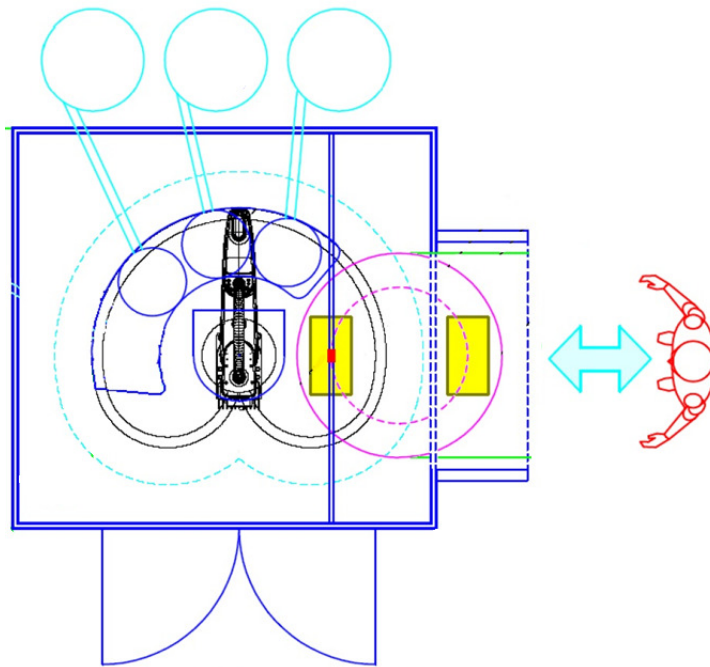


Figura 1. Layout della cella robotizzata.

Dalla vista in pianta è possibile individuare le parti fondamentali che costituiscono la cella robotizzata:

- recinzione perimetrale di sicurezza con porte sensorizzate d'accesso alla cella e barriera a fotocellule per consentire l'accesso alla tavola rotante;
- tavola rotante è a due postazioni, viene comandata dal controllore del robot e permette di portare il prodotto, caricato dall'operatore, all'interno dell'area di lavoro del robot; presenta delle maschere di fissaggio velocemente riconfigurabili per la staffatura dei prodotti durante l'assemblaggio;
- magazzino utensili: permette l'alloggiamento dei diversi organi terminali per il robot ed è direttamente connesso al sistema di alimentazione delle curve;
- sistema di alimentazione delle curve: è costituito da una serie di vibroalimentatori orbitali a gradini; permette il riempimento automatico, in tempo mascherato, dei caricatori degli end-effector depositati a magazzino;
- robot SCARA con relativo controllore e cambio utensile pneumatico;
- end-effector: utensile speciale per il robot realizzato per manipolare ed inserire le curve nel circuito idraulico;
- sistema di visione costituito da una telecamera a matrice e relativo sistema di illuminazione;
- work station per l'utilizzo del software di gestione della cella robotizzata e per la comunicazione con il controllore robot, con il sistema di visione e con il software gestionale dello stabilimento.

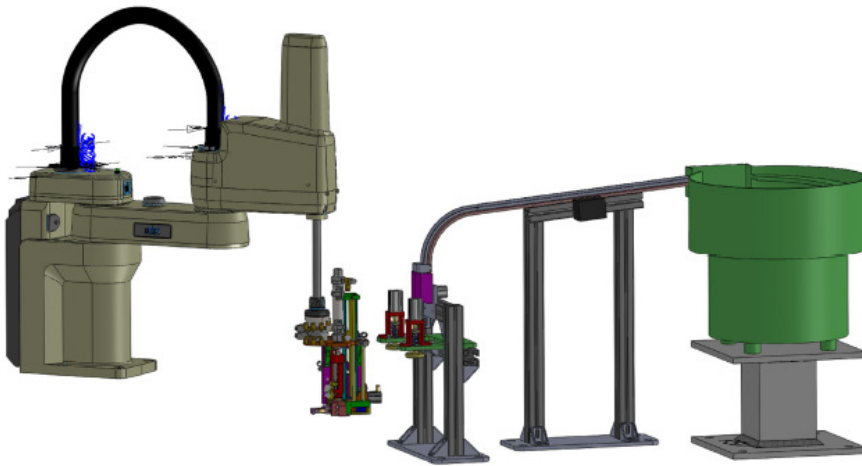


Figura 2. Modello CAD 3D di alcuni sottosistemi dell’impianto.

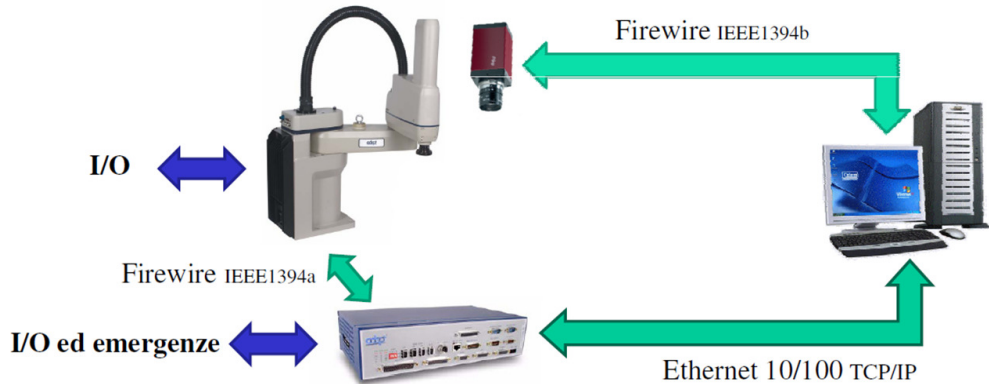


Figura 3. Schema di comunicazione tra i sottosistemi della cella di lavoro.

Progettazione dell’end-effector per il robot

L’end-effector da realizzare, essendo un dispositivo speciale, dedicato all’esecuzione di una particolare applicazione, ha determinate caratteristiche note fin dai primi stadi del progetto. Basandosi sul fatto che il robot non deve realizzare un montaggio delle curve mediante un’operazione di pick-and-place, poiché si è voluto che il percorso che il manipolatore

compie nell'esecuzione dell'assemblaggio fosse tale da rendere minimo il tempo di montaggio, si è posta la specifica iniziale che l'end-effector dovesse possedere un caricatore integrato, da alimentare automaticamente, che consentisse di avere a disposizione lungo tutto il percorso le curve da inserire. Studiando le caratteristiche delle curve si è arrivati a definire la modalità e l'orientazione con cui queste possono essere alimentate nel dispositivo di estremità. Alla luce di questi studi, si è pervenuti a definire la struttura di un caricatore verticale che deve essere riempito dall'alto e al cui interno si trova una guida, di forma opportuna, su cui scorrono per gravità le curve da assemblare, man mano che quelle sottostanti vengono montate sul prodotto. Nella fase di progetto si sono studiati anche i meccanismi necessari perché le curve vengano portate dall'orientamento con cui sono caricate, fino a quello con cui è possibile inserirle. La curva infatti, per le caratteristiche del sistema automatico di alimentazione, entra nel caricatore disponendosi con gli assi dei tubi orizzontali. Per poter essere assemblata sul prodotto, invece, essa dovrà avere gli assi dei tubi verticali e la parte esterna della curvatura rivolta verso l'alto. Inoltre la curva, a causa della sua forma tubolare piegata ad U, è un oggetto difficile da manipolare, si è quindi studiato un sistema in cui la curva è spinta all'interno di una sede, denominata culla, che presenta una cavità della forma della curva. Una volta presa nella culla, la curva è mossa, orientata e trasportata fino al punto di assemblaggio. L'end-effector, inoltre, dovrà svolgere anche la funzione di singolarizzazione delle curve prima di prenderle dalla pila presente nel caricatore, in quanto la scarsa precisione dell'impilaggio non consente di estrarre l'ultima curva dalla pila senza averla preventivamente separata dalle altre.

Perciò l'end-effector realizzato possiede:

- un caricatore in grado di contenere le curve da inserire in un prodotto, il quale è in grado inoltre di potersi ben interfacciare con il sistema di alimentazione automatico;
- un meccanismo di singolarizzazione dell'ultima curva;
- un sistema di presa della curva nella culla;
- un sistema di orientazione che permette di portare la curva con gli assi dei tubi verticali e curvatura verso l'alto;
- un attuatore che permette un movimento di inserimento verticale.



Figura 4. Operazioni di singolarizzazione, orientazione ed inserimento richieste all'end-effector.

Inoltre, accanto a queste caratteristiche che riguardano il suo funzionamento, l'end-effector realizzato è caratterizzato da:

- attuatori compatti, leggeri e veloci che permettono di inserire la curva in un tempo estremamente ridotto;
- permettere di portare la curva da inserire nella giusta orientazione e posizione, rispettando la tolleranza di posizionamento dei tubi nello schema del prodotto;
- avere una struttura che permetta di realizzare un end-effector per ogni tipologia di curva da inserire, cambiando il minor numero possibile di componenti;
- essere dotato di una femmina per il sistema cambio utensile.

Il funzionamento dell'end-effector, come accennato, prevede di far ruotare la culla, contenente la curva, senza che questa si stacchi dalla prima per effetto centrifugo. La rotazione della culla infatti è veloce e il ruolo di contenimento della culla può non essere adeguato a mantenere aderente la curva alla culla nel caso il meccanismo di rotazione non sia ben progettato. Il meccanismo prevede un accoppiamento di forma tra la spina in acciaio alloggiata nella culla e due camme a corsia, poste affacciate e simmetriche rispetto alla culla, per rendere simmetrico il meccanismo ed evitare flessioni, che comprometterebbero il buon orientamento della curva da inserire. Le formule utilizzate per calcolare la traiettoria del centro della spina di acciaio, che costituisce con la camma l'accoppiamento di forma, definiscono i vettori di posizione x_p e y_p i quali, disegnati in un grafico 2D, portano ad ottenere il profilo della camma. Variando i parametri delle formule utilizzate, si è ottenuto un profilo che rende contenuti i valori delle accelerazioni e quindi delle forze d'inerzia.

L'end-effector progettato è stato dapprima realizzato mediante prototipazione rapida, per mezzo di tecnologia FDM, e testato per verificare sincronismo e tempistiche dei meccanismi. Successivamente è stato realizzato un prototipo in materiale metallico, che è stato sottoposto ad analisi mediante lo studio dei filmati dell'assemblaggio a 120 fps. In questo modo è stato possibile ottimizzare le tempistiche ed effettuare alcune modifiche al design, ottenendo una ulteriore riduzione dei tempi di inserimento delle curve.

Per poter eseguire correttamente l'operazione di inserimento di una curva in una determinata locazione è necessario conoscere la posizione e l'orientazione dell'end-effector utilizzato rispetto al sistema di riferimento del robot. A questo scopo si è studiato una metodologia di calibrazione degli organi terminali, ricorrendo ad opportuni utensili, gestita dal software di controllo della cella robotizzata.

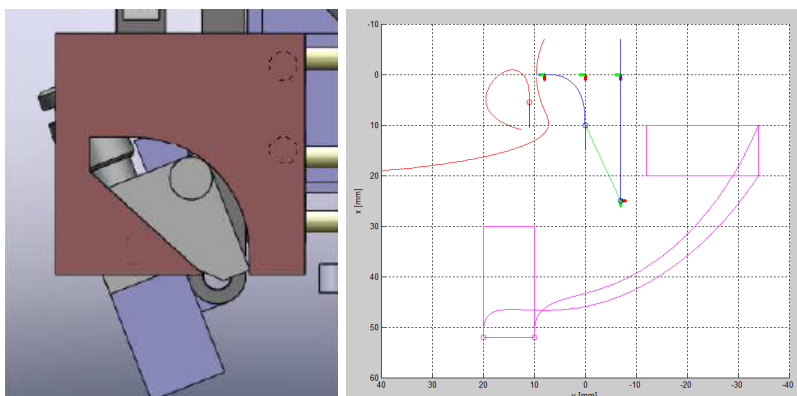


Figura 5. Progetto del meccanismo a camma per la rotazione della curva.

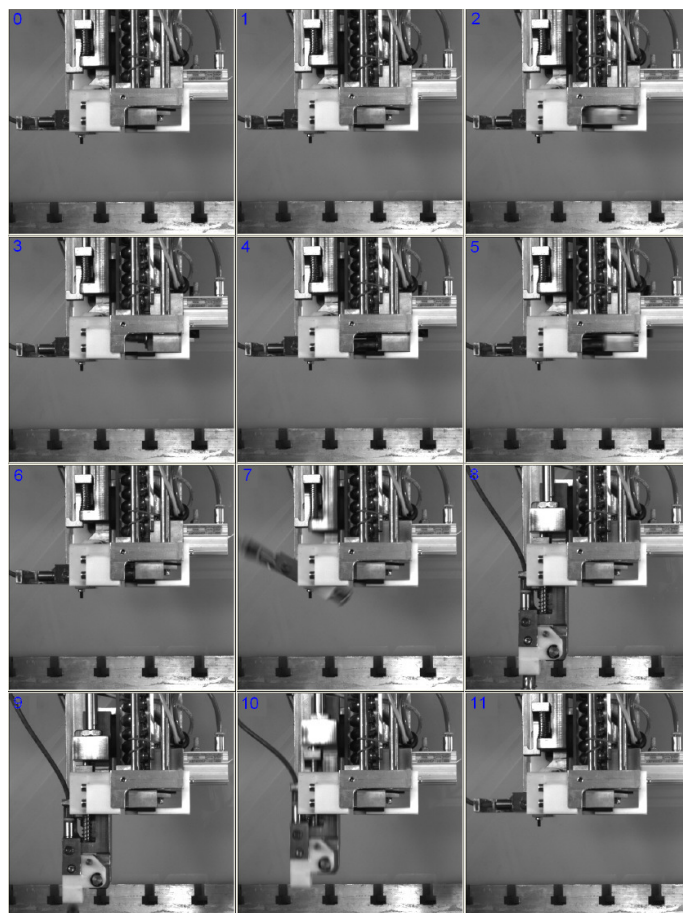


Figura 6. Ottimizzazione delle tempistiche di movimentazione della curva.

Sistema di visione e software di gestione dell'impianto

Il prodotto, una volta posizionato sotto al dispositivo di acquisizione delle immagini, costituito da una telecamera b/n a matrice, viene fotografato all'inizio del ciclo di assemblaggio. Il software di elaborazione delle immagini provvede per prima cosa al riconoscimento del modello dello schema, sulla base delle informazioni relative ai modelli presenti nel database. Nel caso in cui risultino già presenti alcune curve inserite, viene verificato il loro corretto inserimento. Se questa prima verifica iniziale è superata con successo, il software di visione procede ad un'analisi più approfondita di ogni tubo. Vengono rilevate con precisione le coordinate del centro tubo nel sistema di riferimento del robot e valutati eventuali difetti, quali tagli o altre irregolarità nella sommità del tubo che possono compromettere il corretto assemblaggio del prodotto.

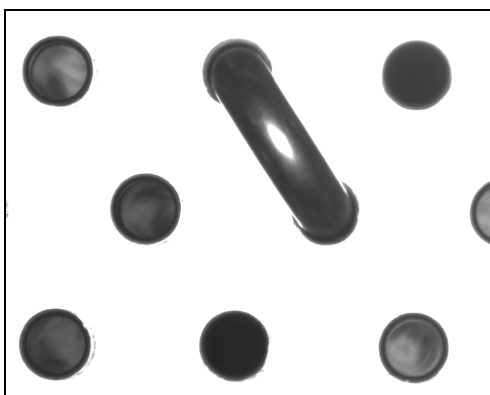


Figura 7. Acquisizione ed elaborazione dell'immagine.

Per definire una corrispondenza tra le coordinate dei centri dei tubi rilevate dal sistema di visione, espresse nel sistema di riferimento bidimensionale dell'immagine, e quelle effettive definite nel sistema di riferimento del robot, si richiede di determinare una trasformazione tra i due sistemi di riferimento. Questa trasformazione viene determinata mediante la procedura di calibrazione della telecamera, sviluppata ad hoc per l'applicazione in esame. L'operazione di calibrazione oltre a definire la corrispondenza tra i due sistemi, permette di compensare tutte le distorsioni prospettiche che caratterizzano questi sistemi, effettuando una mappatura completa del piano di assemblaggio.

Il software di controllo della cella robotizzata è gestito tramite un'interfaccia grafica, che consente di comandare e visualizzare tutte le principali operazioni della cella. A partire dall'interfaccia principale è possibile richiamare altre routine, tra cui il riconoscimento dello schema di montaggio dai file CAD, la gestione degli schemi di lavorazione in corso e il database schemi. Nell'interfaccia principale è possibile seguire direttamente e monitorare l'evoluzione dell'assemblaggio, interagendo con il controllo del robot richiedendo fermi macchina e riattivazioni del ciclo di lavoro. Inoltre è disponibile la visualizzazione di statistiche che rappresentano l'andamento qualitativo delle operazioni di assemblaggio. Si gestiscono e si visualizzano eventuali allarmi associabili a situazione anomala nel cambio utensile, nel robot o nel prodotto. Per quanto riguarda il robot, la sequenza di movimentazione, generata mediante un algoritmo di percorso a tempo minimo di assemblaggio, viene calcolata dal software di gestione, che comunica al controllore del

robot (tramite rete ethernet) le coordinate di movimentazione ed il relativo sottoprogramma da eseguire. Il controllore del robot, oltre a comandare il manipolatore, gestisce fisicamente anche gli input e gli output della cella.

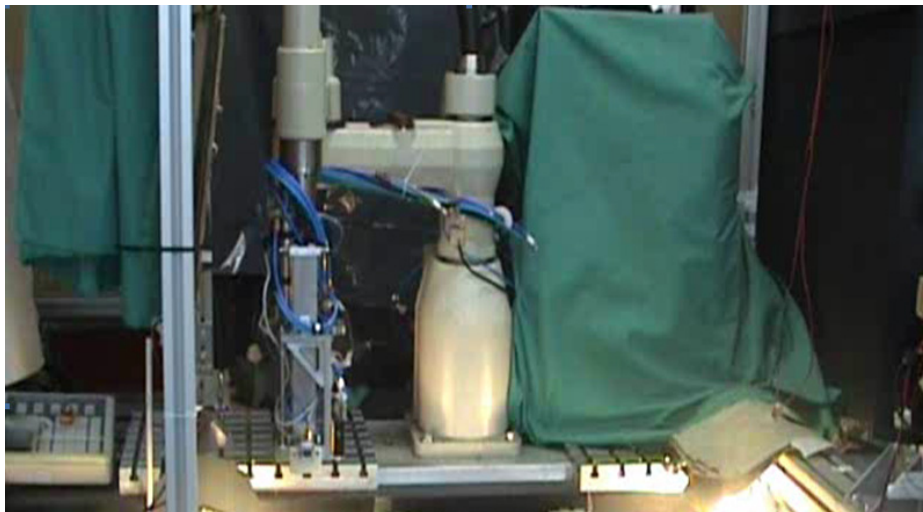


Figura 8. Prototipo della cella robotizzata.

3. CONCLUSIONI

In questo lavoro è stata seguita una procedura metodologica in grado di prendere in considerazione l'insieme delle dinamiche che si presentano nella progettazione esecutiva di un impianto automatico di assemblaggio con alto grado di autonomia e riconfigurabilità. In base ai test svolti e al collaudo effettuato, la cella robotizzata costruita è in grado di assemblare i prodotti con il medesimo tempo-ciclo rispetto alla lavorazione manuale, e necessita di un solo operatore a bordo dell'impianto per il carico/scarico del prodotto e l'alimentazione dei bowl-feeders con le curve da assemblare. È possibile perciò affermare che la cella robotizzata risulta competitiva in termini di tempi e costi di produzione. L'assemblaggio con la cella robotizzata permette inoltre una produzione flessibile, il monitoraggio dei difetti dovuti alle operazioni produttive precedenti, la tracciabilità e la certificazione della qualità dei prodotti ottenuti.

BIBLIOGRAFIA

- [1] K. Feldmann, H. Rottbauer, N. Roth, 1996. "Relevance of assembly in global manufacturing". *CIRP Annals-Manufacturing Technology*, v 45, n 2, pp. 545-552.
- [2] K. Feldmann, S. Slama, 2001. "Highly flexible assembly-Scope and justification". *CIRP Annals-Manufacturing Technology*, v 50, n 2, pp. 489-498.
- [3] W. S. Newman et. al, 2000. "Design lessons for building agile manufacturing systems". *IEEE Transactions on Robotics and Automation*, v 16, n 3, pp. 228-238.
- [4] N. F. Edmondson, A. H. Redford, 2002. "Generic flexible assembly system design". *Assembly Automation*, v 22, n 2, pp. 139-152.

- [5] Reuven Katz, 2007. "Design principles of reconfigurable machines". *International Journal of Advanced Manufacturing Technology*, v 34, n 5-6, pp. 430-439.
- [6] P. G. Ranky, 2003 "Reconfigurable robot tool designs and integration applications". *Industrial Robot: An International Journal*, v 30, n 4, pp. 338-344.
- [7] C. Connolly, 2009. "Machine vision advances and applications". *Assembly Automation*, v 29, n 2, pp. 106-111.
- [8] G. C. Smith, S. S. F. Smith, 2002. "An enhanced genetic algorithm for automated assembly planning". *Robotics and Computer Integrated Manufacturing*, v 18, n 5-6, pp. 355-364.
- [9] G. Rosati, G. Boschetti, A. Biondi, and A. Rossi. "On-line dimensional measurement of small components on the eyeglasses assembly line". *Optics and Lasers in Engineering*, 47(3-4):320–328, 2009.
- [10] G. Rosati, G. Boschetti, A. Biondi and A. Rossi. "Real-time defect detection on highly reflective curved surfaces". *Optics and Lasers in Engineering*, 47(3-4):379–384, 2009.
- [11] V. G. Riello, A. Foroni, G. Rosati, and A. Rossi. "Metodo per l'inserimento di curve nei tubi per fluido di uno scambiatore di calore". Brevetto di invenzione industriale, CCIAA di Milano, Feb-22nd 2010. MI2010A000275.
- [12] V. G. Riello, A. Foroni, S. Cenci, G. Rosati and A. Rossi. "Apparato per l'inserimento di curve nei tubi per fluido di uno scambiatore di calore". Brevetto di invenzione industriale, CCIAA di Milano, Feb-22nd 2010. MI2010A000267.

MECHANICAL DESIGN OF BIMATERIAL HELICAL SPRINGS WITH CIRCULAR CROSS-SECTION

Eugenio Dragoni

Dipartimento di Scienze e Metodi dell'Ingegneria,

Università di Modena e Reggio Emilia, Italia

E-mail: eugenio.dragoni@unimore.it

William J. Bagaria

Aerospace Engineering Department,

United States Naval Academy

E-mail: gfwjb207@verizon.net

Abstract. *This paper presents an exact mechanical model for helical springs with circular cross-section formed by a inner elastic core encased in an outer annulus of dissimilar elastic properties. Closed-form equations are developed for stresses and deflection in the spring undergoing either bending or axial end loads. For both loading conditions, the model takes into account the stress concentrations arising in the cross-section due to curvature of the spring axis. The disclosed equations are specialized for bimaterial springs with polymer core and thin nanometal cladding, a solution reflecting a unique technology recently brought into the market by a leading polymer manufacturer. In this special case, the cladding behaves as an efficient thin-walled tube under torsion with the soft core preventing the danger of wall instability. A design procedure is exemplified, showing that this construction leads to lighter and smaller springs than all-metal or all-polymer counterparts.*

Keywords: *mechanical design, helical springs, circular section, bimaterial construction.*

1. INTRODUCTION

Analysis and design equations are available for helical springs in many shapes and materials. For the most common solid sections and homogeneous materials the classical theory developed by Wahl [1] applies. The advantages of hollow-section helical springs were first pointed out by Bagaria [2] and his results were included by Spinella and Dragoni [3] in a theory for hollow shape memory springs. In [4] Gobbi and Mastinu developed a general theory for tubular composite springs with anisotropic material behaviour.

Recently, a proprietary material process [5] has been disclosed which strengthens and stiffens polymer parts by means of a thin cladding (0.1-0.3 mm) of nanometals deposited on the outer surface (Figure 1). The remarkable mechanical properties of typical cladding alloys are listed in Table 1. The new technology can be applied to helical springs by coating a polymer blank produced either by rapid prototyping or, for simpler geometries, by high-volume manufacturing techniques such as extrusion. The high-performance nanometal cladding works very efficiently in round-section helical springs because it is ideally placed to resist torsion. Lightweight and compact springs, well suited for light- and

medium-duty service (biomedical devices, electrical equipment, automotive interiors), result from this advantage.

Table 1: Mechanical properties of NiFe nanometals compared with spring steels

Property	Unit	Nano NiFe (alloy A)	Nano NiFe (alloy B)	High-strength steel	Stainless steel
Tensile strength	MPa	1 710	1 440	2 300	1 000
Yield strength	MPa	1 180	950	2 200	870
Yield shear strength	MPa	680	548	1 270	502
Young's modulus	GPa	147	141	206	195
Shear modulus	GPa	56	54	79	75
Elongation to failure	%	5	8	6	6
Mass density	kg/dm ³	8.6	8.6	7.9	7.9

No theory exists for helical springs based on this hybrid construction. By ignoring the polymer core, the theory for hollow springs [3] can be applied but the contribution of the core to strength and stiffness is lost, at times with appreciable errors. This paper develops an exact theory for bimaterial helical springs with circular cross-section undergoing either end axial forces or end bending moments (Figure 2). Equations are presented for stress and deflection which generalize the corresponding expressions of the classical theory for homogeneous springs.

The disclosed general theory is specialized for bimaterial springs with polymer core and thin nanometal cladding, for which case the equations become very simple. It is to be noted that for a hollow spring, the spring cross-section, if too thin, can collapse [3]. Adding a polymer core to a thin cladding will prevent this collapse. By means of a numerical example, a design procedure is presented for this particular spring construction. The procedure starts from specifications on load and deflection and supplies manufacturing spring data (material, section and coil diameters, number of coils) together with spring properties (solid length, mass). Length and mass of the nanometal-cladded spring are compared with the corresponding properties of a conventional steel spring subjected to the same design requirements.

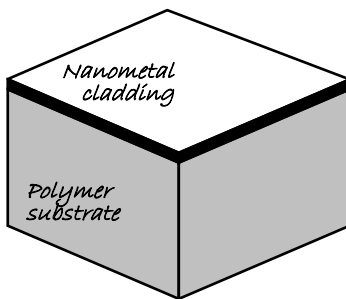


Figure 1. Hybrid system formed by nanometal cladding on polymer substrate

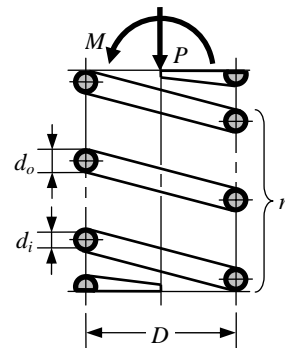


Figure 2. Section of bimaterial helical spring with design variables and load conditions.

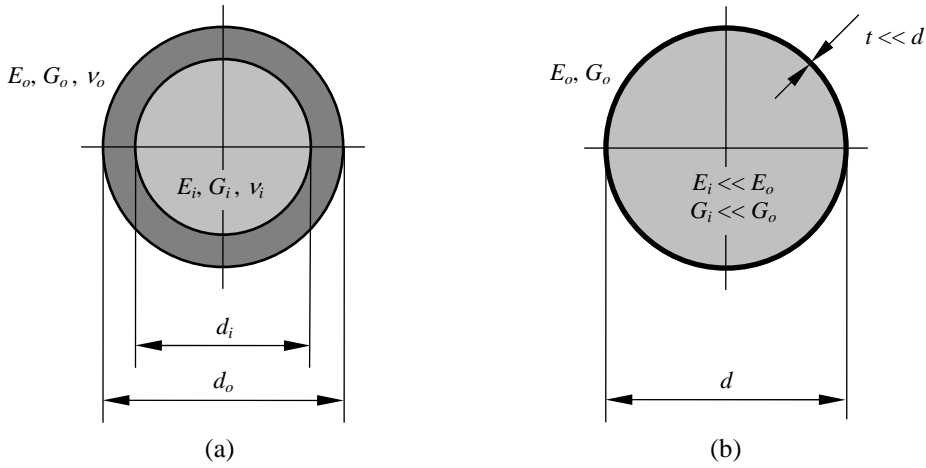


Figure 3. Dimensions and elastic properties of the bimaterial cross-section: general case (a); limit case for very thin cladding (b).

2. EQUIVALENT SECTION PROPERTIES

The general properties and dimensions of the bimaterial cross-section are given in Figure 3a, with pedix “i” referring to the inner core and pedix “o” referring to the outer cladding. Since core and cladding are connected together, the torsional and the bending stiffness of the section are calculated as the sum of the individual stiffnesses of the two parts [6]. The equivalent torsional stiffness becomes:

$$\langle GJ_p \rangle = \frac{\pi}{32} \left\{ G_i d_i^4 + G_o (d_o^4 - d_i^4) \right\} \quad (1)$$

The equivalent bending stiffness is given by:

$$\langle EJ \rangle = \frac{\pi}{64} \left\{ E_i d_i^4 + E_o (d_o^4 - d_i^4) \right\} \quad (2)$$

For core and cladding with equal Poisson’s ratio, ν , combination of (1) and (2) gives:

$$\langle GJ_p \rangle = \frac{\langle EJ \rangle}{(1 + \nu)} \quad (3)$$

3. SPRING UNDER END BENDING MOMENT

Shear Stresses

The cross-sections of the spring in the plane of bending (Figure 2) develop torsional shear stresses. The nominal shear stress at the outside of the cladding is obtained by

dividing the applied moment by the torsional stiffness (1) and multiplying by the product $G_o d_o / 2$ [6] between the shear modulus and the outer radius of the cladding:

$$\tau_{M \text{ nom}} = \frac{M}{\langle GJ_p \rangle} G_o \frac{d_o}{2} = \frac{16MG_o d_o}{\pi \left\{ G_i d_i^4 + G_o (d_o^4 - d_i^4) \right\}} \quad (4)$$

Because of the curvature of the spring, the shear stress is greater at the points of the cladding surface facing the spring axis. The maximum stress at this point is calculated by applying to (4) a curvature stress concentration factor:

$$\tau_{M \text{ max}} = K_M \tau_{M \text{ nom}} = K_M \frac{16MG_o d_o}{\pi \left\{ G_i d_i^4 + G_o (d_o^4 - d_i^4) \right\}} \quad (5)$$

with K_M given by:

$$K_M = \frac{4c(1-b^2 + \chi b^2) - 1 + b^4 - \chi b^4}{4(c-1)(1-b^2 + \chi b^2)} \quad (6)$$

where $c = D/d_o$ is the spring index, $b = d_i/d_o$ is the cladding aspect ratio and $\chi = G_i/G_o$ is the core to cladding stiffness ratio. Factor K_M is plotted in Figure 4 as a function of c for the limit cases of homogeneous solid section ($b = 0$, solid line) and infinitely thin hollow section ($b = 1$, $\chi = 0$, dashed line). For $0 < b < 1$ and $0 < \chi < 1$, factor K_M assumes values falling between the two curves in Figure 4.

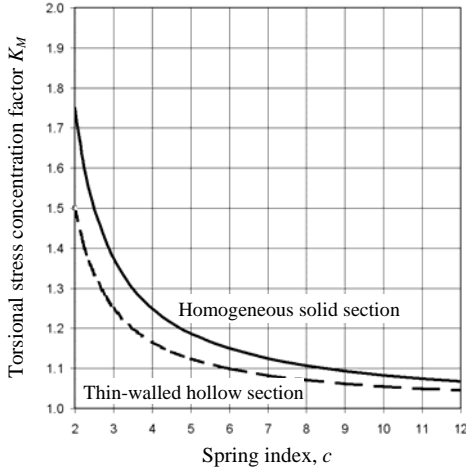


Figure 4. Plot of the curvature stress concentration factor K_M for solid and hollow sections.

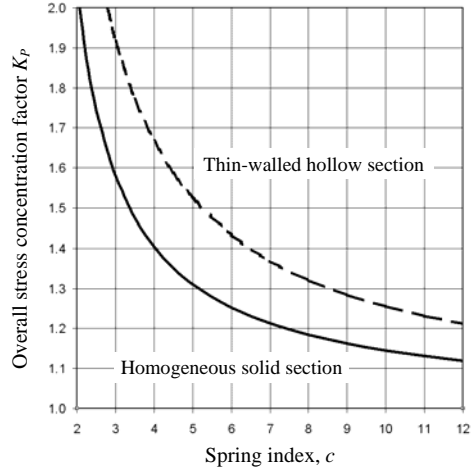


Figure 5. Plot of the curvature stress concentration factors K_p for solid and hollow sections.

Normal Stresses

The cross-sections at right angle with the plane of bending develop bending normal stresses. The maximum bending stress at the outside of the cladding is obtained by dividing the applied moment by the bending stiffness (2) and multiplying by the product $E_o d_o / 2$ [6] between the Young's modulus and the outer radius of the cladding:

$$\sigma_{M_{max}} = \frac{M}{\langle EJ \rangle} E_o \frac{d_o}{2} = \frac{32ME_o d_o}{\pi \{E_i d_i^4 + E_o (d_o^4 - d_i^4)\}} \quad (7)$$

There is no curvature correction for the bending stress because the plane of curvature of the spring is at right angle with the plane of bending.

Deflection

Timoshenko [7] gives the end rotation α of a homogeneous helical spring under bending moment in terms of torsional and bending stiffnesses of the cross-section. For the bimaterial spring, the end rotation is obtained from Timoshenko's expression by referring to the equivalent stiffnesses:

$$\alpha = \frac{\pi MDn}{2} \left(\frac{1}{\langle GJ_p \rangle} + \frac{1}{\langle EJ \rangle} \right) \quad (8)$$

Assuming the same Poisson's ratio, ν , for core and cladding, from (1)-(3) equation (8) becomes

$$\alpha = \frac{32MDn(2+\nu)}{E_i d_i^4 + E_o (d_o^4 - d_i^4)} \quad (9)$$

4. SPRING UNDER END AXIAL FORCE

Shear Stresses

The axial end load P in Figure 2 produces shear stresses in the cross-section due to torsional and transverse shear contributions. The torsional contribution can be calculated from (5) by letting $M = PD/2$. Following [8], the maximum transverse shear stress in the cladding can be written as:

$$\tau'_{P_{max}} = K'_P \frac{8PDG_o d_o}{\pi \{G_i d_i^4 + G_o (d_o^4 - d_i^4)\}} \quad (10)$$

Assuming the same Poisson's ratio, ν , for core and cladding [8], the curvature factor K'_P is given by:

$$K'_p = \frac{1}{c} \cdot \frac{(1-b^2)\{1+2\nu+b^2(3+2\nu)\} + \chi\{1+2\nu-2b^2+b^4(3+2\nu)\}}{2(1+\nu)\{1-b^2+\chi(1+b^2)\}} \quad (11)$$

Letting $M = PD/2$ in (5) and adding the resulting expression to (10), the maximum total shear stress in the cladding is obtained as:

$$\tau_{p_{max}} = K_p \frac{8PDG_o d_o}{\pi \{G_i d_i^4 + G_o (d_o^4 - d_i^4)\}} \quad (12)$$

where $K_p = K_M + K'_p$. Using (6) and (11), K_p expands to:

$$K_p = \frac{4c(1-b^2 + \chi b^2) - 1 + b^4 - \chi b^4}{4(c-1)(1-b^2 + \chi b^2)} + \frac{1}{c} \cdot \frac{(1-b^2)\{1+2\nu+b^2(3+2\nu)\} + \chi\{1+2\nu-2b^2+b^4(3+2\nu)\}}{2(1+\nu)\{1-b^2+\chi(1+b^2)\}} \quad (13)$$

Factor K_p is plotted in Figure 5 as a function of c for the limit cases of homogeneous solid section ($b = 0$, solid line) and infinitely thin hollow section ($b = 1$, $\chi = 0$, dashed line). Again, for $0 < b < 1$ and $0 < \chi < 1$, the values of K_p fall between the two curves in Figure 5.

Deflection

The spring under load P in Figure 2 can be likened to a straight rod of length $L = \pi nD$, where n is the number of active coils. The angular twist β of the end sections under torsional moment $T = PD/2$ is $\beta = TL / \langle GJ_p \rangle$. Combining these expressions with (1) gives the axial deflection $\delta = \beta D/2$ as:

$$\delta = \frac{8PD^3 n}{G_i d_i^4 + G_o (d_o^4 - d_i^4)} \quad (14)$$

5. DESIGN EQUATIONS FOR NANOMETAL CLADDING

The particular case of a bimaterial section with polymer core and nanometal cladding (Table 1) is illustrated in Figure 3b. The section is identified by mean diameter, d , and thickness, t , of the cladding, with $t \ll d$. The elastic properties of the polymer core are also much lower than those of the cladding. Under these assumptions, the response of the spring can be obtained from the general equations developed in Sections 3 and 4 by letting

$d_o \approx d_i \approx d$, $d_o - d_i \approx 2t$, $d_o^4 - d_i^4 \approx 8td^3$, $G_i \approx 0$ and $E_i \approx 0$. The following subsections show the final results.

Spring under End Bending Moment

Expression (5) for the maximum shear stress becomes:

$$\tau_{M \max} = \bar{K}_M \frac{16M}{\pi d^3} \cdot \left(\frac{0.125}{t/d} \right) \quad (15)$$

where \bar{K}_M is the limit value for $b = d_i/d_o = 1$ and $\chi = G_i/G_o = 0$ of the curvature stress concentration factor (6), which becomes (dashed curve in Figure 4):

$$\bar{K}_M = \frac{2c-1}{2(c-1)} \quad (16)$$

Likewise, the maximum bending stress (7) becomes:

$$\sigma_{M \max} = \frac{32M}{\pi d^3} \cdot \left(\frac{0.125}{t/d} \right) \quad (17)$$

Finally, the end rotation (9) simplifies to:

$$\alpha = \frac{32MDn(2+\nu)}{E_o d^4} \cdot \left(\frac{0.125}{t/d} \right) \quad (18)$$

Spring under End Axial Force

The maximum shear stress (12) becomes:

$$\tau_{P \max} = \bar{K}_P \frac{8PD}{\pi d^3} \cdot \left(\frac{0.125}{t/d} \right) \quad (19)$$

with \bar{K}_P , limit case for $b = d_i/d_o = 1$, $\chi = G_i/G_o = 0$ and $\nu = 0.3$ of the curvature stress concentration factor (13), given by (dashed curve in Figure 5):

$$\bar{K}_P = \frac{2c-1}{2(c-1)} + \frac{2}{c} \quad (20)$$

The deflection (14) can be written:

$$\delta = \frac{8PD^3n}{G_o d^4} \cdot \left(\frac{0.125}{t/d} \right) \quad (21)$$

It is seen that equations (15), (17), (18), (19) and (21), giving the properties of the spring, all contain function ψ defined as:

$$\psi = \frac{0.125}{t/d} \quad (22)$$

Function ψ is plotted in Figure 6 against the ratio t/d in the range 0.04–0.2 .

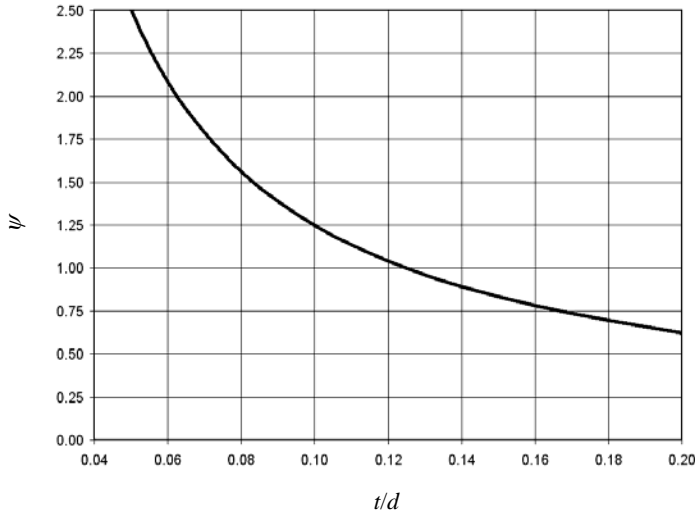


Figure 6: Plot of function ψ used in equations (15)–(21).

Design Example

Using the equations developed in Section 5.2, let's design a bimaterial spring with nylon (PA66) core and nanometal cladding subjected to a maximum axial force $P = 150$ N and maximum deflection $\bar{\delta} = 15$ mm. Choosing nanoalloy A from Table 1, the following material properties are identified for the cladding: yield strength $\bar{\tau} = 680$ MPa, shear modulus $G_o = 56\,000$ MPa, mass density $\rho_o = 8.6$ kg/dm³ = 8.6×10^{-3} g/mm³. Assuming a spring index $c = 7$, the stress concentration factor $\bar{K}_p = 1.37$ is calculated from (20). Assuming $t = 0.30$ mm and letting $\tau_{Pmax} = \bar{\tau}$, the mean diameter of the cladding is calculated from (19) as:

$$d = \bar{K}_p \frac{8P}{\pi \bar{\tau}} \cdot \left(\frac{D}{d} \right) \cdot \left(\frac{0.125}{t} \right) = 1.37 \times \frac{8 \times 150}{\pi \times 680} \times 7 \times \frac{0.125}{0.30} \approx 2.2 \text{ mm} \quad (23)$$

giving a mean coil diameter $D = c d = 7 \times 2.2 = 15.4 \text{ mm}$ and an outside coil diameter $D_o = D + d = 15.4 + 2.2 = 17.6 \text{ mm}$. Finally, for $\delta = \bar{\delta}$, equation (21) gives the number of active coils:

$$n = \frac{\bar{\delta} G_o d^3}{8 P D^3} \cdot \left(\frac{t}{0.125} \right) = \frac{15 \times 56\,000 \times 2.2^3}{8 \times 150 \times 15.4^3} \times \frac{0.30}{0.125} \approx 4.9 \quad (24)$$

Assuming $n_0 = 2$ dead turns, the total number of turns becomes $n_T = n + n_0 = 4.9 + 2 = 6.9$, giving a solid length $H = n_T d = 6.9 \times 2.2 = 15.2 \text{ mm}$ and a total “wire” length $L = \pi D n_T = \pi \times 15.4 \times 6.9 \approx 334 \text{ mm}$. The volume of the core is $V_i = L \pi (d - t)^2 / 4 = 334 \times \pi \times (2.2 - 0.3)^2 / 4 = 947 \text{ mm}^3$, which, for a mass density $\rho_i = 1.1 \times 10^{-3} \text{ g/mm}^3$, gives the mass $m_i = \rho_i V_i = 1.1 \times 10^{-3} \times 947 \approx 1.0 \text{ g}$. The cladding has a volume $V_o = L \pi d t = 334 \times \pi \times 2.2 \times 0.3 = 693 \text{ mm}^3$, with a mass $m_o = \rho_o V_o = 8.6 \times 10^{-3} \times 693 \approx 6.0 \text{ g}$. The total mass of the bimaterial spring is $m = m_i + m_o = 1.0 + 6.0 = 7.0 \text{ g}$.

Using textbook equations [9], a solid steel spring with the same yield stress as the cladding, subjected to the same load, undergoing the same deflection and wound with the same spring index as the bimaterial spring is readily designed. The following characteristics are found: wire diameter $d = 2.2 \text{ mm}$, mean coil diameter $D = 15.4 \text{ mm}$, outside coil diameter $D_o = 17.6 \text{ mm}$, number of active coils $n = 6.3$, solid height $H = 18.3 \text{ mm}$ and total mass $m = 11.9 \text{ g}$ (for $n_0 = 2$ dead turns). Likewise, a glass-reinforced solid nylon spring with yield strength $\bar{\tau} = 120 \text{ MPa}$, shear modulus $G_i = 5\,770 \text{ MPa}$, mass density $\rho_i = 1.6 \times 10^{-3} \text{ g/mm}^3$ and spring index $c = 4$ (the value giving the lowest mass) would have the following characteristics: wire diameter $d = 4.2 \text{ mm}$, mean coil diameter $D = 16.8 \text{ mm}$, outside coil diameter $D_o = 21.0 \text{ mm}$, number of active coils $n = 4.7$, solid height $H = 28.1 \text{ mm}$ and total mass $m = 7.9 \text{ g}$ (for $n_0 = 2$ dead turns).

6. DISCUSSION

It is easily verified that for $d_i = 0$ the equations for stresses (including curvature factors) and deflections developed in Sections 3 and 4 reduce to the well known results for the homogeneous solid spring [1]. Similarly, when $G_i = 0$, the general equations coincide with those developed in [3] for hollow springs.

Figure 4 shows that the curvature stress concentration factor under pure torsion, K_M , is always higher for the solid spring (solid line) than for the thin-walled hollow spring (dashed line). The ratio between the dashed curve and the solid curve decreases with the spring index c , achieving the minimum value 0.87 for $c = 2$. By contrast, when the overall stress concentration factor K_p is considered (Figure 5), this behaviour is reversed. Now the overall factor of the thin-walled hollow spring (dashed line) exceeds that of the solid spring (solid line). The ratio between the dashed curve and the solid curve increases for decreasing c and stabilizes at the upper value 1.22 for $c \leq 3$. For truly bimaterial sections defined by any $b = d_i/d_o$ ratio between 0 and 1 and stiffness ratio $\chi = G_i/G_o$ greater than

zero, the curvature factors lie between the dashed and solid curves in Figures 4 and 5. The exact values can be calculated using equations (6) for K_M and (13) for K_P .

Function ψ defined in (22) and plotted in Figure 6 has an important physical meaning. Disregarding the slight difference in the curvature factors K_M and K_P between solid section and thin hollow section, function ψ represents the ratio between any property (stress or deflection) of the thin-cladded polymer spring and the corresponding property of the all-nanometal solid spring. When $t/d = 0.125$, the thin-cladded spring and the all-nanometal spring have equal properties ($\psi = 1$). This means that a cladding as thin as $t = 0.3$ mm with average section diameter $d = 0.3/0.125 = 2.4$ mm would produce the same maximum stress and deflection of an all-nanometal solid section with the same diameter d . This example gives an idea of the efficiency of the cladding in this torsional application.

The efficiency of the cladding is further confirmed by the design example in Section 5.3, with the cladded nylon spring being 41% lighter than the steel spring and 14% lighter than the glass-reinforced nylon spring. The bimaterial spring is also shorter than the steel spring (17% lesser solid length) and of the nylon spring (less 46%). Although the outside diameter of the cladded spring is equal to the steel spring's, it is 16% lower than the nylon spring's. These values correspond to spring configurations that minimize the mass for given load and deflection. If the spring parameters of homogeneous steel and nylon springs are changed (for instance by changing the spring index c), some properties of the spring can slightly improve (solid length or outside diameter) but only at great expense of the mass.

7. CONCLUSIONS

An exact mechanical model for helical springs with bimaterial circular cross-section has been developed. The model leads to closed-form equations for stresses and deflection when the spring undergoes either bending or axial end loads. In both loading cases, the theory incorporates the stress concentrations arising in the cross-section due to curvature of the spring axis. The disclosed equations are specialized for bimaterial springs with polymer core and thin nanometal cladding, a new technology recently introduced by a leading polymer manufacturer. In this special case, the cladding behaves as an efficient thin-walled tube under torsion with the soft core preventing wall instability that would occur in a hollow spring. A numerical example shows the benefits of this construction in terms of reduced mass and overall dimensions with respect to all-metal or all-polymer springs.

REFERENCES

- [1] A. M. Wahl, *Mechanical Springs*, McGraw-Hill, New York, 1963.
- [2] W.J. Bagaria, "Stress analysis of light weight closely coiled hollow helical springs", Annual summary of the United States Naval Academy, www.usna.edu/AcResearch/2000SummaryPDFs/DivEngrWeapAerospaceEngrDept.pdf, 2000.
- [3] I. Spinella, E. Dragoni, "Analysis and design of hollow helical springs for shape memory actuators", *J. Intel. Mat. Sys. Struct.*, 21, 185-199, 2009.
- [4] M. Gobbi, G. Mastinu, "On the optimal design of Composite Material tubular Helical spring", *Meccanica*, 36, 525-553, 2001.

- [5] M.R. Day, “Nanometal-polymer hybrid”, *Advanced Materials and Processes*, 166, 25-27, 2008.
- [6] D. Gay, S.V. Hoa, S.W. Tsai, *Composite materials: design and applications*, CRC Press, 2003.
- [7] S. Timoshenko, D.H. Young, *Elements of Strength of Materials*, D. VanNostrand, Princeton, 1962.
- [8] K.H. Lo, H.D. Conway, “Effect of change of Poisson’s ratio on the bending of a multi-layered circular cylinder”, *Int. J. Mech. Sci.*, 16, 757-767, 1974.
- [9] J. Shigley, C. Mischke, R. Budynas, *Mechanical engineering design*, McGraw-Hill, New York, 2003.

APPENDIX (Notation)

b	=	Aspect ratio of cladding ($= d_i/d_o$)
c	=	spring index ($= D/d_o$)
d	=	mean diameter of cladding layer (also outer diameter of solid-section spring)
d_i, d_o	=	inner and outer diameters of cladding
D	=	mean diameter of spring coil
E_i, E_o	=	Young’s modulus of core and cladding
$\langle EJ \rangle$	=	equivalent bending stiffness of bimaterial section
G_i, G_o	=	shear modulus of core and cladding
$\langle GJ_p \rangle$	=	equivalent torsional stiffness of bimaterial section
H	=	solid length of spring
K_M	=	curvature stress concentration factor under moment end load
\bar{K}_M	=	value of K_M for thin rigid cladding on soft core
K'_P	=	curvature stress concentration factor for transverse shear under axial end load
K_P	=	overall curvature stress concentration factor under axial end load
\bar{K}_P	=	value of K_P for thin rigid cladding on soft core
L	=	length of active coils
L_T	=	total length of spring wire (including dead coils)
m_i, m_o	=	mass of core and cladding
m	=	total mass of spring
M	=	moment end load
n, n_0	=	number of active and dead coils
n_T	=	total number of spring coils
P	=	axial end load
t	=	thickness of cladding layer
T	=	torsional moment on spring section under axial end load

V_i, V_o	=	volume of core and cladding
α	=	angular deflection of spring under moment end load
β	=	angular twist between end sections of active coil length
δ	=	axial deflection of spring under axial end load
$\bar{\delta}$	=	design deflection of spring under axial end load
ν	=	Poisson's ratio common to core and cladding
ν_i, ν_o	=	Poisson's ratio of core and cladding
ρ_i, ρ_o	=	mass density of core and cladding
$\sigma_{M \max}$	=	maximum bending stress in cladding due to moment end load
$\bar{\tau}$	=	allowable shear stress of cladding
$\tau_{M \text{ nom}}$	=	nominal shear stress in cladding due to moment end load
$\tau_{M \max}$	=	maximum shear stress in cladding due to moment end load
$\tau'_{P \max}$	=	maximum transverse shear stress in cladding due to axial end load
$\tau_{P \max}$	=	maximum overall shear stress in cladding due to axial end load
χ	=	core to cladding stiffness ratio ($= G_i / G_o$)
ψ	=	dimensionless function
i, o	=	pedices referring to inside and outside surfaces of cladding

CONSIDERAZIONI SUI CRITERI DI RESISTENZA

Gianluca Medri

Dipartimento di Ingegneria Industriale,

Università di Parma

E-mail: gianluca.medri@unipr.it

Sommario. *Si presentano considerazioni sull'impiego dei tradizionali Criteri di Resistenza nella progettazione e verifica degli organi strutturali di macchine e manufatti. In riferimento ad una rappresentazione non-convenzionale dello spazio operativo (in termini di tensioni principali) si puntualizzano caratteristiche di efficienza e difetti di affidabilità dei Criteri.*

Keywords: *Criteri di resistenza. Verifica di resistenza. Resistenza dei materiali.*

1. INTRODUZIONE

Uno dei tanti problemi che si pongono al progettista meccanico nella fase dell'*embodiment* è l'individuazione, prima qualitativa e poi quantitativa, dell'indicatore del livello di sollecitazione della struttura. Questo indicatore deve risultare inferiore, in confronti deterministici (coefficienti di sicurezza) o stocastici (analisi probabilistiche di resistenza), alla resistenza del materiale nel punto considerato nel pezzo reale, generalmente individuata da σ_R (tensione normale di collasso: rottura, snervamento o cedimento strutturale in senso lato).

Il campo tensionale è identificato dalle tre tensioni principali " $\sigma_1, \sigma_2, \sigma_3$ " (convenzionalmente σ_1 ha il valore più elevato in senso algebrico etc.) e dalle relative direzioni nello spazio. Molto rozzamente (e anche scorrettamente, da un punto di vista rigorosamente teorico) si può affermare che il problema di cui si tratta nasce dal fatto che, per semplicità tecnica e scientifica (sostanzialmente per evitare una spaventosa messe di prove sperimentali), la resistenza del materiale è individuata da un solo "numero" (il valore di σ_R) mentre il livello di sollecitazione nel corpo è quantificato da sei numeri e per trasformare sei numeri in un unico indicatore occorre elaborare delle teorie.

Quindi, mentre per identificare la generica resistenza dei materiali strutturali è sufficiente fare prove esaurienti di rottura (usualmente abbastanza semplici), l'indicatore della sollecitazione è ottenuto ipotizzando le modalità di rottura (**teoria o ipotesi di rottura**) del materiale in esame e riformulando il modello fisico in termini di tensioni, fino ad ottenere un "numero" (**σ ideale**) da confrontare con σ_R (R_s o R_m), generando un **criterio di resistenza**.

Oramai le teorie di rottura, e i relativi criteri di resistenza, sono talmente entrate sotto la pelle di generazioni d'ingegneri e di ricercatori che spesso si tende a banalizzare la scelta della " σ ideale" senza più tenere conto delle approssimazioni e imprecisioni, implicite in ognuno di questi criteri, che sono ben note agli studiosi ma spesso molto meno agli utilizzatori (perché dimenticate o non considerate interessanti dal punto di vista pratico).

I codici di calcolo FEM offrono solitamente le mappe degli indicatori, come opzioni di

output, almeno del Criterio del Lavoro di Distorsione. E' un gran sollievo per il calcolatore, ma può allontanare e offuscare la parte più conscia della progettazione, perché alcuni materiali e alcune situazioni di carico non sono bene rappresentate da questo particolare modello. Inoltre la tecnologia informatica *Knowledge Based Design* impiega pacchetti applicativi dove i criteri di analisi strutturale sono talmente stratificati "in profondità" che risulta difficile un'analisi critica dei risultati ottenuti dal loro impiego.

Sono notissimi i diagrammi che confrontano i campi di sicurezza ottenibili applicando vari criteri di resistenza in un piano convenzionale σ_1 - σ_2 (in letteratura sono attribuiti a Westergaard), ipotizzando nulla la terza tensione principale (σ_3 ?). Questi diagrammi, di valenza prevalentemente didattica, mostrano che i vari criteri di resistenza si differenziano nel risultato pratico delle analisi strutturali permettendo, così, di giungere alla conclusione che alcuni di essi sono meno "conservativi" di altri (Fig. 1) per l'ampiezza relativa delle zone di resistenza. In questa sede si tralasciano le "imperfezioni" della rappresentazione (solo uno dei semispazi identificato dalla bisettrice dei quadranti I e III è di interesse - l'altro è simmetrico - e non è rispettato l'ordine delle tensioni principali) per concentrare l'attenzione sui Criteri.

In Fig. 1 (per comportamento duttile) sono rappresentati: il Criterio della Massima Tensione Normale (quadrato a trattini, $R_s = \sigma_1$ o $-R_s = \sigma_3$), il Criterio della Massima Tensione Tangenziale (esagono "inclinato" a tratti, $R_s = \sigma_1 - \sigma_3$) e il Criterio della Massima Energia di Distorsione (ellisse a tratti e punti, $R_s^2 = [(\sigma_1 - \sigma_2)^2 + (\sigma_1 - \sigma_3)^2 + (\sigma_2 - \sigma_3)^2]/2$).

E' sottaciuto che, in certe zone dello spazio " σ_1 - σ_2 - σ_3 ", alcuni criteri presentano carenze sensibili di modellazione, o prevedono comportamenti non accettabili perfino "a buon senso". Questo fatto è certamente più grave di un ottimismo quantitativo di modellazione, perché non esiste coefficiente di sicurezza che possa proteggere da un modello non fisico.

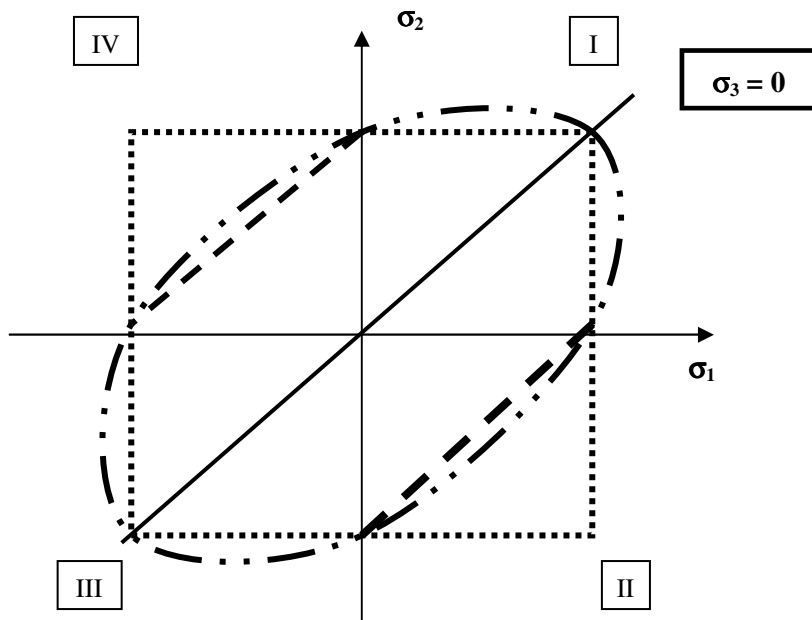


Figura 1. Campi di resistenza

Il problema, infatti, nasce dalla constatazione che nei diagrammi del tipo di quelli di Fig. 1 il campo di resistenza è rappresentato come **chiuso** e questo può spingere qualche utilizzatore a pensare che ognuno di questi criteri preveda **limiti oggettivi** sicuri e ben precisi a livello tensionale per ogni tipo di sollecitazione applicabile. Ma, ovviamente, questa è solo un'impressione grafica: in alcune configurazioni dello stato tensionale, che nei diagrammi di Fig. 1 non possono essere rappresentate, non esiste un limite oggettivo alla sollecitazione ma rimane solo il limite sulla σ ideale (**artificiosamente**) creato dall'ipotesi di rottura specifica.

Si potrebbe pensare che, in fondo, è solo questo che ci si aspetta da un criterio di resistenza. Invece no! Il criterio di resistenza deve approssimare il comportamento reale del materiale al meglio, e non viceversa. Non è "fisico" che in certe situazioni il modello del materiale resista indefinitamente all'incremento di sollecitazione; inoltre ciò che si può accettare in compressione non è, usualmente, altrettanto accettabile in trazione.

La rappresentazione in un piano genericamente identificato come " σ_1 - σ_2 " è, quindi, incapace di evidenziare questi "buchi", anche per l'equivoca etichettatura delle tensioni principali (la terza tensione è assunta nulla), e la sua vera utilità è confinata al campo meramente didattico (e di basso livello). L'ipotesi di stato piano di tensione, che è alla base di queste rappresentazioni, è applicabile con buoni o ottimi risultati a molte situazioni strutturali critiche ma non a tutte!

Purtroppo, spesso il progettista non ha il tempo di approfondire il problema proposto, la cui analisi in realtà non è molto complessa, ma richiede solo un attimo di attenzione ai modelli e alle ipotesi alla base della progettazione. Per questo l'autore ritiene che valga la pena tracciarne uno sviluppo a scopo puramente esemplificativo e con esplicita esclusione di ogni riferimento a sollecitazioni affaticanti. L'analisi che segue è limitata a modelli per materiali isotropi.

2. LIMITI DI RESISTENZA NEL PIANO σ_1 - σ_3

Per meglio valutare i modelli di rottura, e relativi criteri di resistenza, è più funzionale il piano σ_1 - σ_3 , **con le tensioni principali ordinate secondo il pedice** come usuale nella Meccanica dei Continui. Per evitare di analizzare il problema in 3D, eventualmente utile solo in un caso ben specifico, conviene escludere σ_2 . Rispetto ai diagrammi classici questo è solo un accorgimento tecnico (e, d'altra parte, non vuole essere niente di più) che però permette una visione più oggettiva della situazione tensionale. In questa rappresentazione, la situazione presa come riferimento nei diagrammi di Westergaard (stato piano di tensione) è contenuta nel II quadrante del diagramma (σ_1 positiva e σ_3 negativa) e nei suoi confini. In tutti i diagrammi possibili di questo tipo, la zona di resistenza è compresa tra la bisettrice del I e III quadrante (nel diagramma proposto questa retta separa il campo NON FISICO – in alto a sinistra – dal campo FISICO – in basso a destra - e rappresenta il luogo degli stati idrostatici di tensione dove $\sigma_1=\sigma_2=\sigma_3$) e le linee limite dei singoli criteri.

La verifica di resistenza nel diagramma proposto può impiegare il concetto della retta di carico: dall'origine degli assi (stato a tensioni nulle) si traccia la retta con coefficiente angolare pari al rapporto tra σ_1 e σ_3 nello stato tensionale di interesse. L'intersezione tra la retta di carico e il confine di resistenza individua la situazione di collasso per il sistema di sollecitazioni ipotizzato e permette di valutare il coefficiente di sicurezza come rapporto tra una tensione di collasso e la sua omologa di riferimento in esercizio.

Ragionando dal punto di vista puramente (e intuitivamente) fisico, ipotizzando per ogni

materiale l'esistenza di una tensione (normale) "limite di rottura per distacco" σ^* , **situazione di rottura fragile**, e di una tensione (tangenziale) "limite di rottura per scorrimento" τ^* , **situazione di rottura duttile**, si può costruire il diagramma di Fig. 2, dove sono riportati tre casi notevoli di coppie di valori di σ^* e τ^* .

La linea retta verticale a tratti rappresenta il limite di resistenza a trazione su σ_1 (σ^*) mentre le linee a tratti inclinate a 45° individuano il limite di resistenza $\sigma_1 - \sigma_3 = 2\tau^*$. La bisettrice del II quadrante rappresenta stati tensionali in cui $\sigma_1 = -\sigma_3$ (sollecitazioni di taglio puro nel piano delle due tensioni), mentre la parte positiva dell'asse delle σ_1 "contiene" le sollecitazioni di trazione uniassiale (non solo quelle, ovviamente).

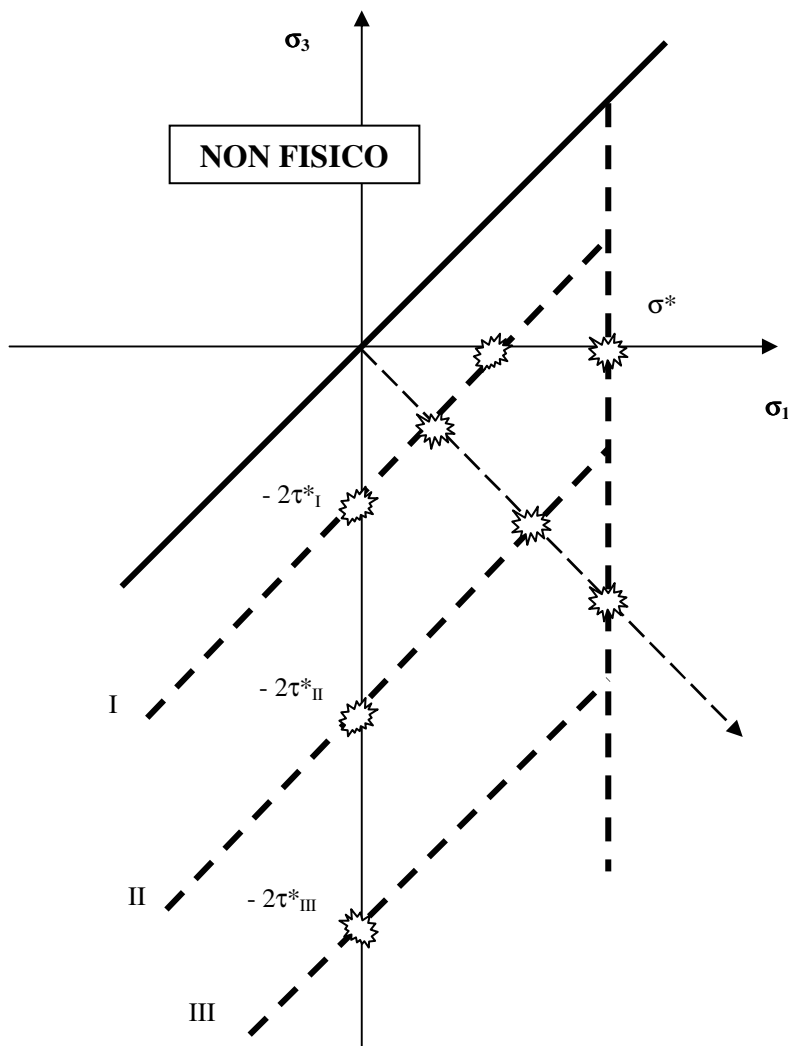


Figura 2. Rotture per distacco e per scorrimento nel piano σ_1 - σ_3 .

Sono state evidenziate le situazioni di collasso prevedibili sulla base di questa semplicistica approssimazione alla rottura dei materiali: il materiale generico può esibire comportamenti duttili sia a trazione uniassiale sia a torsione (I caso), comportamento duttile a torsione e fragile a trazione uniassiale (II caso) e comportamento fragile sia a trazione uniassiale sia a torsione (III caso). A compressione uniassiale (contenuta nella parte negativa dell'asse delle ordinate) il comportamento risulterebbe sempre duttile (a parte possibili rotture fragili per effetto Poisson, con piano di separazione parallelo alla direzione di compressione). Per stati tensionali tendenzialmente idrostatici di compressione la resistenza tende ad aumentare senza limiti. Già in base a questo semplice esempio si vedono comportamenti del materiale non usualmente contemplati dal progettista medio.

Nella versione più semplice del **Criterio di Resistenza della Massima Tensione Normale** (indicatore di pericolosità σ_1) la rappresentazione si ricava da Fig. 2 cancellando le rette inclinate a 45° del comportamento duttile: rimane il solo limite $\sigma^* = \sigma_R$ (Fig. 3). In questo caso, la resistenza a compressione uniassiale risulterebbe infinita se non s'introducesse anche un limite sulla σ_3 .

Questo limite può essere giustificato dall'esistenza di una τ^* (come riportato in figura, linea a tratti inclinata a 45°), per quanto alta sia, oppure legato alle deformazioni trasversali in allungamento per effetto Poisson (e quindi ricollegato alla σ^* ad esempio con il desueto Criterio della Massima Deformazione, come indicato in Fig. 3).

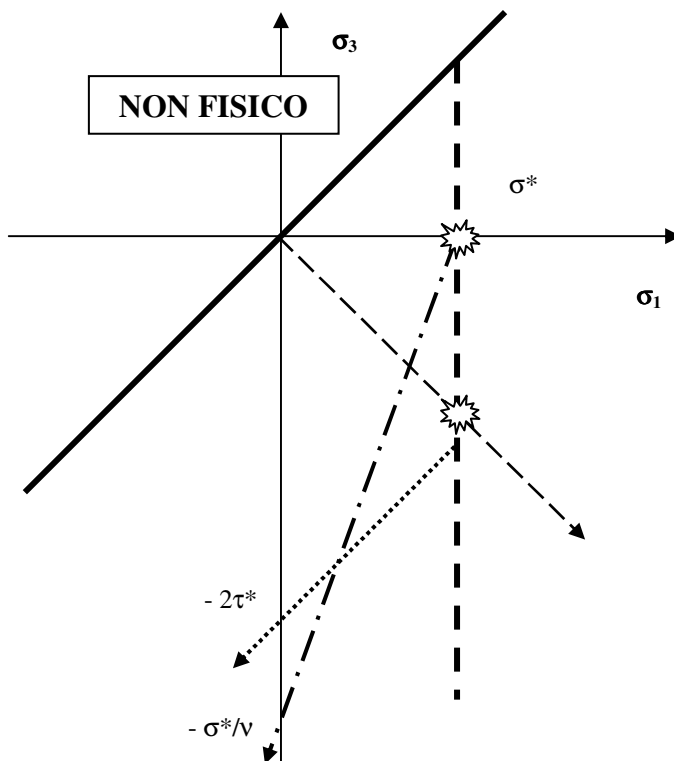


Figura 3. Criterio della Massima Tensione Normale.

In ogni caso, questo diagramma rivela l'incerta funzionalità di questo criterio quando vi siano tensioni di compressione non trascurabili. Il risultato non è tanto eclatante in quanto è già noto a sufficienza (o almeno dovrebbe esserlo) e per questa ragione, usualmente, l'applicazione di questo criterio è riservata ai materiali “fragili” che hanno problemi sostanziali molto più grossi nelle zone in trazione.

Passando al **Criterio di Tresca** (limitazione della τ_{MAX} , per comportamento duttile), la rappresentazione proposta in Fig. 4 mostra alcuni problemi di modellazione del comportamento fisico del materiale per stati tensionali di trazione a basso valore della τ_{MAX} .

La linea tratteggiata che individua il criterio non limita il valore delle tensioni genericamente massime di trazione o compressione, modellando un comportamento francamente poco fisico, soprattutto dal lato della trazione, nei quadranti I e III.

Per quanto riguarda, infine, il **Criterio di Von Mises** (o del Lavoro di Distorsione), in Fig. 5 le rette I e II, tratteggiate parallele e inclinate di 45° rispetto agli assi, sono rispettivamente i limiti della zona “resistente” I) per σ_2 coincidente con σ_1 o con σ_3 , II) per $\sigma_2 = (\sigma_1 + \sigma_3)/2$.

La curva a tratti nel II quadrante rappresenta il confine di resistenza per tensione piana ($\sigma_2 = 0$); anche i segmenti AB e A'B' contengono le situazioni di collasso in caso di tensione piana (rispettivamente per $\sigma_3 = 0$ e per $\sigma_1 = 0$).

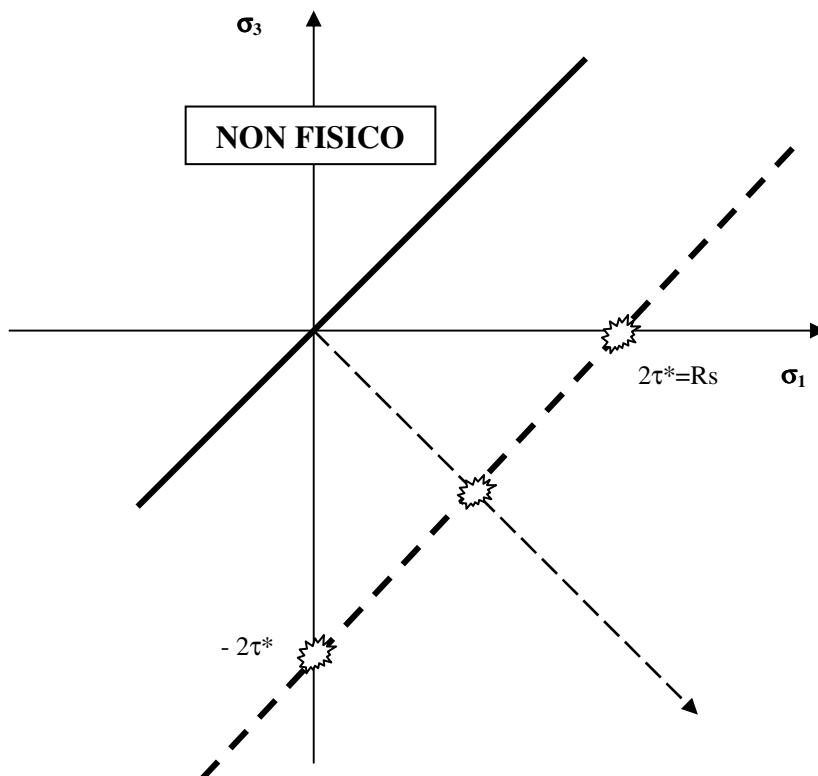


Figura 4. Criterio della Massima Tensione Tangenziale.

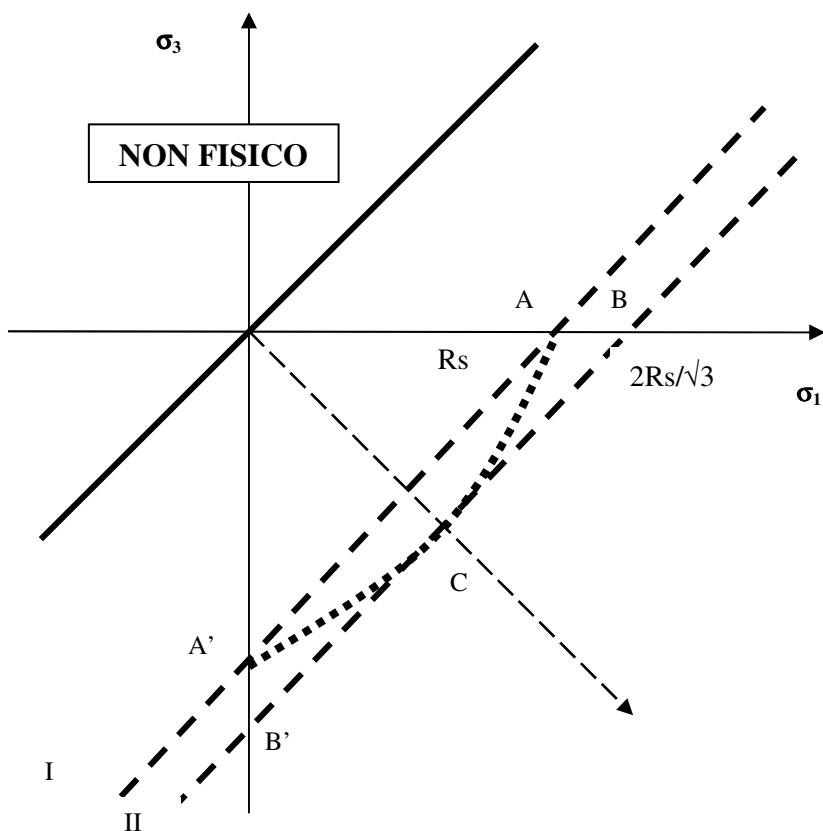


Figura 5. Criterio del Lavoro di Distorsione.

Per questo criterio l'uso della retta di carico è più complesso che nei casi precedenti, perché l'intersezione con la linea limite è determinata in riferimento anche al valore della σ_2 . In ogni caso, però, i confini di resistenza per i possibili valori di σ_2 (che per definizione è la tensione principale con valore intermedio) sono compresi nella porzione di spazio delimitata dalle rette I e II. La tensione di collasso a trazione uniassiale, identificata con R_s , è individuata dall'intersezione della retta I con l'asse delle ascisse (sul quale $0 = \sigma_3 \leq \sigma_2 \leq \sigma_1$); l'intersezione della retta II con l'asse delle ascisse è in $\sigma_1 = 2R_s/\sqrt{3} = 1.155R_s$ e si riferisce ad uno stato tensionale caratterizzato da $\sigma_2 = \sigma_1/2$ e $\sigma_3 = 0$.

I riscontri sperimentali riportati in letteratura confermano la validità del Criterio per sollecitazioni di tensione piana. Nella rappresentazione classica dei dati sperimentali riportata in Fig. 6 ($S_C = R_s$), ripresa da [1] per due tipi di acciaio (cerchi), due leghe leggere (quadrati) e una ghisa (triangolo), si può notare che:

1. I dati sperimentali sopra la bisettrice del I e III quadrante (linea a tratto e punto aggiunta in questa sede) sono "inutili" per la caratterizzazione perché la situazione è perfettamente simmetrica dal punto di vista tensionale, con lo scambio delle tensioni principali (nel I quadrante σ_1 di Fig. 6 diventa σ_2 - tensione principale

- intermedia dello stato tensionale – e viceversa).
2. In conseguenza di quanto precedentemente detto i dati sperimentali del I quadrante sono orientativamente riferiti (a parte leggere oscillazioni) a una tensione variabile $\sigma_{2(1)} = 0 \div \sigma_{1(2)}$, una fissa ($\sigma_3 = 0$) e una leggermente variabile ($\sigma_{1(2)} \approx \sigma_R$ intesa come sollecitazione di collasso – R_s per duttile e R_m per fragile). Questa situazione nel diagramma di Fig. 5 è individuata dal segmento AB e **non descrive stati tensionali nel I quadrante di detto diagramma.**
 3. Il II quadrante di Fig. 6 si “riversa” nel II quadrante di Fig. 5 con la trasformazione di σ_2 nella terza tensione principale. Si notano le situazioni di taglio puro (sulla bisettrice del quadrante, linea a tratti aggiunta in questa sede) e il diverso comportamento dei materiali duttili e della ghisa che verificano la congruità dei criteri rappresentati per confronto.
 4. I dati sperimentali riportati nel diagramma σ_1 - σ_3 indicano una carenza qualitativa di descrizione del comportamento dei materiali (giustificabile anche da difficoltà tecnologiche) non parimenti evidenziata dal diagramma classico (si veda Fig. 6bis, derivata dalla 6 con il cambio di assi, con i quadranti I e III desolatamente vuoti!).

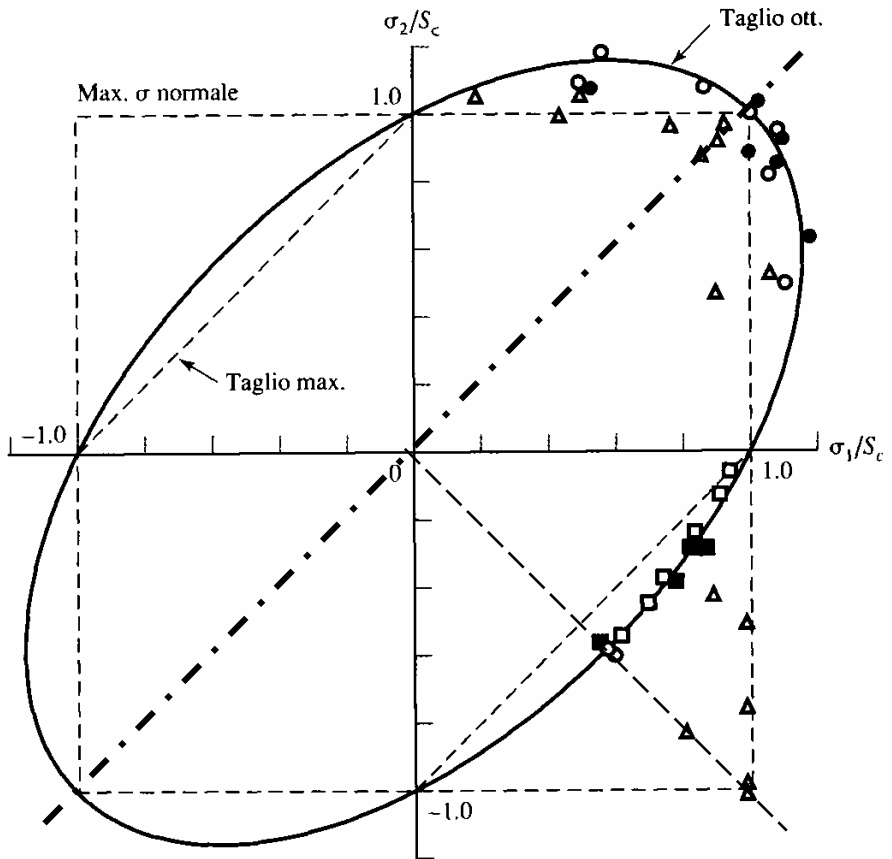


Figura 6. Dati sperimentali [1] (riproduzione da Fig. 7.11 pag. 257 di [2]).

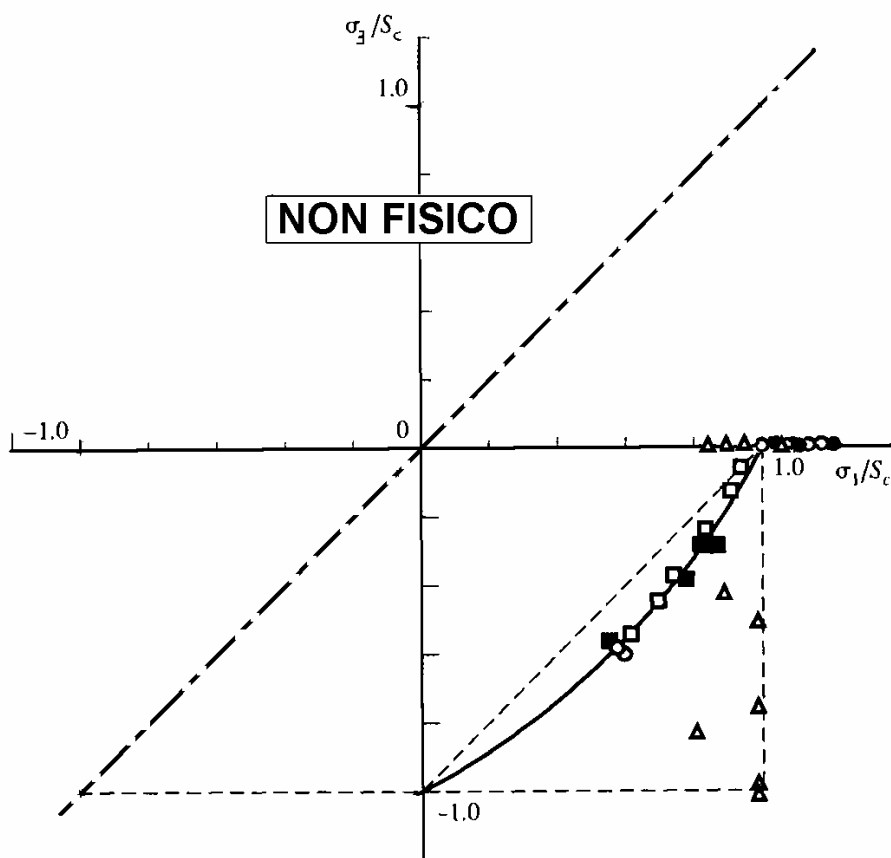


Figura 6bis. Dati sperimentali da Fig. 6 su diagramma σ_1 - σ_3 normalizzato a R_s .

In conclusione, risalta una specifica carenza di modellazione dei comportamenti per stati tensionali a sensibile componente idrostatica. In particolare, per tensioni principali positive e livello distorsivo calante (ad esempio in stato di deformazione piana o campi di tensioni termiche) il modello comportamentale duttile potrebbe perdere di precisione qualitativa (e quindi quantitativa) perché il materiale tende ad assumere comportamenti fragili se lo snervamento è impedito. Ne consegue una sostanziale sottostima della pericolosità da parte dell'indicatore del Lavoro di Distorsione che consiglia un'analisi incrociata dei dati riferiti alle tensioni ideali con quelli relativi alle tensioni principali, specialmente per i risultati FEM.

Ad una prima sommaria verifica con la rappresentazione proposta, il **Criterio di Mohr** (Fig. 7, in versione “fragile”, ma applicabile anche a comportamenti duttili) risulta pragmaticamente il meglio configurato di quelli presentati (fatto da tempo riconosciuto).

Nella Fig. 7 sono riportate due diverse applicazioni del Criterio di Mohr:

- limiti di resistenza identificati dalla spezzata a tratti ABCDE (il solo tratto CD è verificato sperimentalmente!).

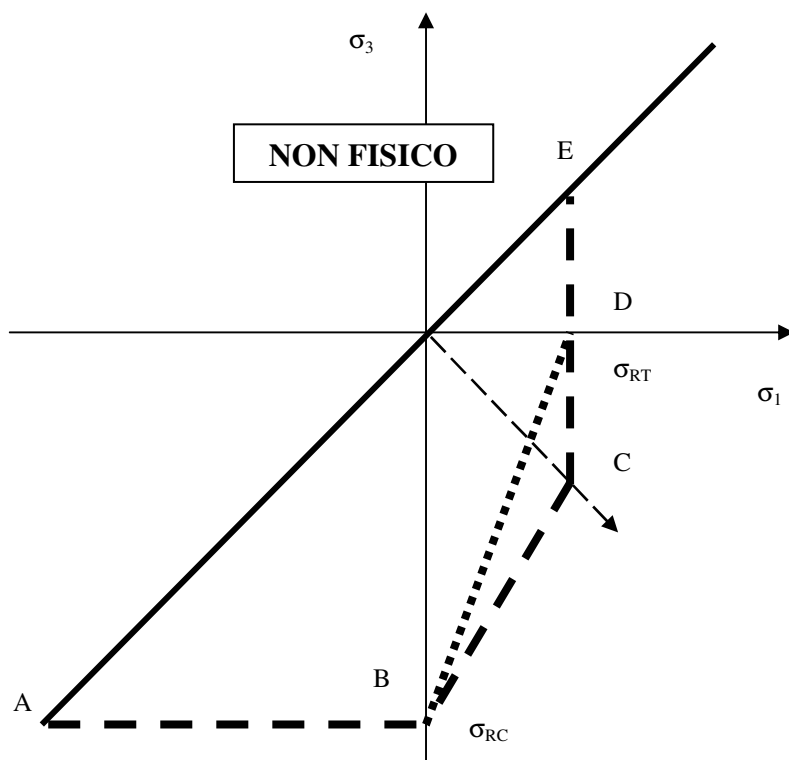


Figura 7. Criterio di Mohr per materiali a comportamento fragile.

- limiti di resistenza identificati dalla spezzata a tratti ABDE (σ_{RC} potrebbe coincidere con $-\sigma^*/\nu$ in riferimento al Criterio della Massima Deformazione).

E' abbastanza ovvio che un criterio fenomenologico (quale, in sostanza, è quello di Mohr) si adatti meglio a modellare un comportamento fisico, con una decente precisione sia dal punto di vista qualitativo che quantitativo qualora l'ampiezza della sperimentazione effettuata per la individuazione della formulazione della "legge" (non solo, quindi, per la determinazione delle costanti del materiale) sia congrua con il campo di applicazione della stessa. La caratteristica fenomenologica e marcatamente grafica (che sembra complicare i calcoli matematici) portano a scartare questo criterio in campo applicativo.

In verità, un criterio di tipo puramente fenomenologico, derivato concettualmente dal Criterio di Mohr si potrebbe basare sul diagramma di Fig. 8, dove il pedice R identifica situazioni di collasso di frattura per comportamenti fragili e di snervamento per comportamenti duttili dei materiali e tutte le caratteristiche di resistenza indicate vanno determinate sperimentalmente in relazione agli effetti di σ_2 (in condizione restrittiva). Le linee a pallini identificano restrizioni conservative del campo di resistenza.

Si ricorda per completezza un criterio proposto (ma non ancora convalidato) per il campo dei materiali elastomerici (che hanno comportamento tendenzialmente fragile): il **Criterio della Massima Tensione Idrostatica** (di trazione, vedi Fig. 9); esso fu "spinto" dal fatto che, in alcune situazioni e per campi tensionali nel I quadrante, si sono riscontrate

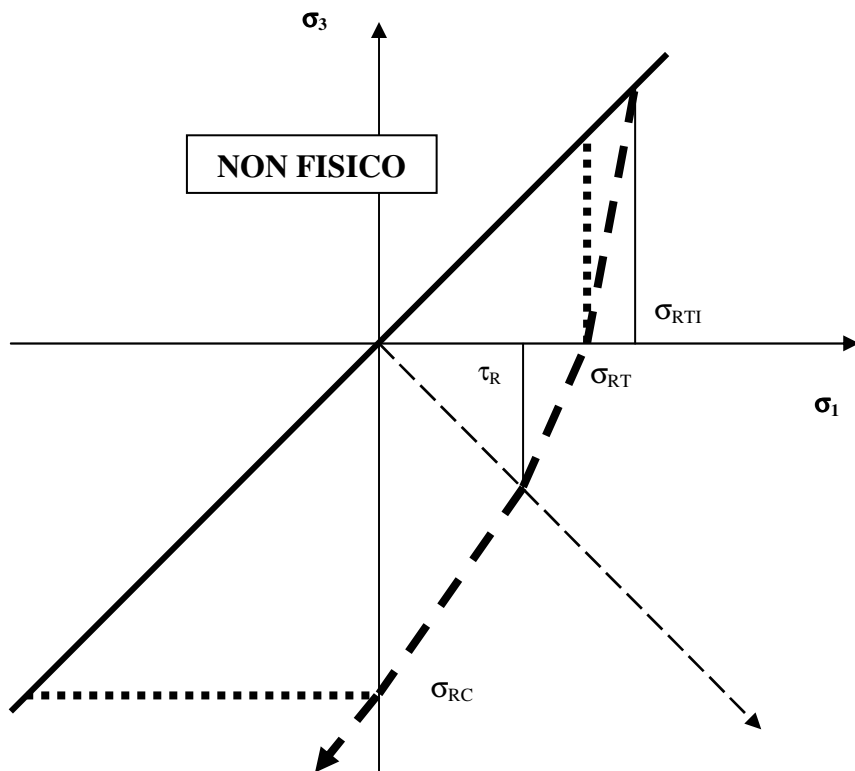


Figura 8. Criterio fenomenologico.

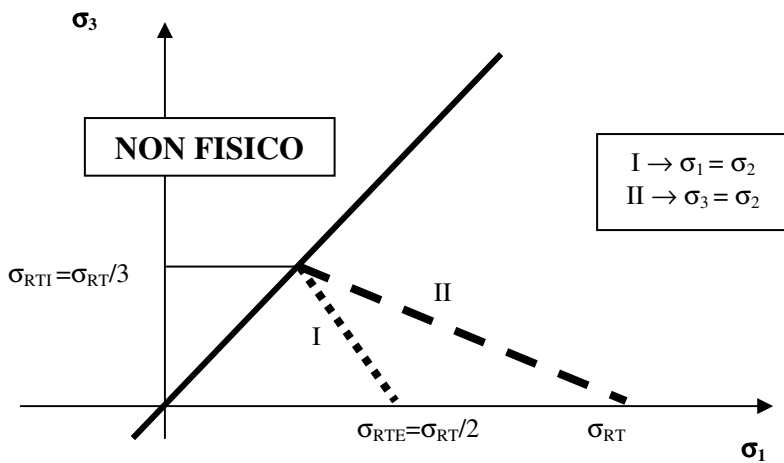


Figura 9. Criterio della Massima Tensione Idrostatica.

rottture non giustificabili con i Criteri tradizionali. Nelle materie plastiche o elastomeriche si ottengono facilmente stati di trazione idrostatica (o quasi idrostatica), ad esempio con il *pocker-chip test*.

Nel *pocker-chip test* un sottile strato di materiale ad elevata cedevolezza (e coefficiente di Poisson vicino a 0.5) viene inglobato tra due elementi di materiale rigido e sollecitato a trazione perpendicolarmente al suo piano medio (generando uno stato di deformazione nulla nel piano dello strato). Questo permette l'identificazione della tensione di rottura in situazioni quasi-idrostatiche (σ_{RTI}), e il Criterio della Massima Tensione Idrostatica assume come indicatore di collasso la $\sigma_{TI} = (\sigma_1 + \sigma_2 + \sigma_3)/3$. La σ_{RTI} si ricava anche da prove di trazione uniassiale ($\sigma_{RTI} = \sigma_{RT}/3$) o equibiassiale ($\sigma_{RTI} = 2\sigma_{RTE}/3$). I segmenti che rappresentano i confini di resistenza sono compresi tra i tratti I e II al variare di σ_2 tra σ_1 e σ_3 .

Questo criterio non appare adatto ai materiali metallici: infatti nei dati sperimentali riportati in Fig. 6 all'aumentare della tensione media non diminuisce la tensione massima di collasso.

3. SCELTA DEL CRITERIO DI RESISTENZA

La dizione “scelta del criterio di resistenza” individua due azioni apparentemente simili ma molto diverse per metodica e scopi immediati:

1. scelta del criterio in sede di progetto/verifica strutturale avendo a disposizione esclusivamente dati di letteratura sul materiale e conoscendo lo stato tensionale della struttura sotto analisi con buona precisione (almeno qualitativa);
2. identificazione del criterio più adeguato per un dato materiale, per il quale sia possibile effettuare una caratterizzazione sperimentale dal punto di vista della resistenza meccanica.

Nel primo caso si devono individuare i campi più sollecitati del corpo e i quadranti interessati dei diagrammi nel piano σ_1 - σ_3 . Note le caratteristiche dichiarate del materiale (fragile o duttile) e i dati sui carichi di rottura che vengono, poi, riportati sul diagramma è possibile valutare quale criterio risulta più adatto (cioè più preciso o, in alternativa, più conservativo) alla situazione tensionale, che può essere rappresentata con rette di carico.

Nel secondo caso deve essere pianificata un'adeguata campagna di prove per riuscire a discriminare, in base ai risultati degli esperimenti, tra i criteri utilizzabili. E' evidente dai diagrammi mostrati che nel II quadrante del piano σ_1 - σ_3 alcuni criteri di resistenza sono praticamente equivalenti, mentre la zona critica è la parte “fisica” del I quadrante. Ciò significa che le semplici prove tecnologiche standard (a trazione uniassiale, a compressione uniassiale, a flessione e a taglio o torsione) non sono sufficienti a discriminare tra i modelli offerti perché non mettono in luce cosa accada al materiale in questo quadrante. Il modello di resistenza meccanica nel quadrante I è, quindi, costruito in base ad una “estrapolazione” e non con “interpolazioni”.

La σ_3 (di trazione) “fa la differenza”: incrementare il solo valore di prova della σ_2 è praticamente inutile per i fini di cui si parla, perché sposta il punto di collasso da A a B (e poi da B ad A) di Fig. 5 per il criterio di Von Mises e non genera alcun effetto per gli altri criteri tradizionali (a parte il Criterio della Massima Tensione Idrostatica). In pratica un effetto misurabile e statisticamente certo della σ_2 (a $\sigma_3 \leq 0$) permette solo di convalidare modelli di Von Mises in riferimento al II quadrante. Per i materiali metallici, prove a σ_3 positiva (per approfondire e affinare il modello di rottura) sono ritenute giustificabili solo

quando la σ_1 nel componente superi una soglia di sicurezza definita in relazione a R_s e contemporaneamente l'indicatore relativo al criterio ne sia sensibilmente inferiore.

4. CONCLUSIONI

Ogni verifica di resistenza effettuata con un qualsivoglia criterio di resistenza è affetta da imprecisioni quantitative di natura concettuale legate al criterio stesso e non generalmente identificabili con facilità o per *default*, oltretutto dalle imprecisioni esclusivamente quantitative caratteristiche di tutte le modellazioni.

Valori elevati degli indicatori di sollecitazione (cioè vicini all'indicatore di resistenza del materiale) sono sicuramente pericolosi, qualunque sia il criterio. Valori bassi degli indicatori nei criteri di Tresca e Von Mises quando sia $\sigma_3 > 0$, non rappresentano una situazione conservativa in sé e per sé e suggeriscono un controllo con il Criterio della Massima Tensione. **I criteri disponibili attualmente sono verificati sperimentalmente solo per stato piano di tensione.**

Inoltre, la scelta del criterio da usare non può essere automatica (un certo materiale non richiede necessariamente “quel” determinato criterio) né univoca (a seconda degli stati tensionali alcuni criteri sono più accurati, e quindi sicuri, di altri).

BIBLIOGRAFIA

- [1] Shigley J. E., Mischke C. R., Budynas R. G., 2005. Progetto e Costruzione di Macchine, McGraw-Hill Ed., Milano.
- [2] Bowling N.E., 1999. Mechanical Behavior of Materials, 2nd Ed. Prentice Hall, Englewood Cliffs, NJ.

PROGETTAZIONE A FATICA TERMOMECCANICA DI GETTI IN LEGA D'ALLUMINIO PER APPLICAZIONI MOTORISTICHE

Gianni Nicoletto

Dipartimento di Ingegneria Industriale

Università degli Studi di Parma

43100 Parma

E-mail: gianni.nicoletto@unipr.it

DEDICA

Questo lavoro è dedicato al prof. Gustavo Favretti sotto la cui tutela ho iniziato la mia carriera e che ricordo per la grande umanità e per il rigore nell'adempire al ruolo accademico.

***Sommario.** Viene presentata una innovativa metodologia di progettazione di componenti motore ottenuti per colata di lega d'alluminio sollecitati termomeccanicamente. La particolarità della metodologia sta nella forte integrazione degli aspetti progettuali/strutturali, metallurgici e tecnologici e nell'uso intensivo di strumenti di calcolo e di simulazione di processo.*

***Parole chiave:** progettazione a fatica, teste motore, elementi finiti, leghe Al-Si, getti*

1. INTRODUZIONE

La tendenza nello sviluppo dei motori moderni evidenzia incrementi di potenza a parità di cilindrata e contenimento o riduzione del peso unitario per ridurre i consumi (i.e. crescente potenza specifica data dal rapporto potenza/massa o kW/kg). Si osserva per questo un uso diffuso di leghe d'alluminio anche per la realizzazione degli elementi strutturali più critici quali teste e basamenti motore, vedi Fig. 1. L'adozione delle leghe d'alluminio nella costruzione di parti motore e trasmissione al posto dei tradizionali materiali ferrosi contribuisce a riduzioni di peso del 50% con un risparmio anche di 70kg per unità.

Siccome le condizioni di lavoro di parti motore sono molto severe (i.e. elevate temperature e sollecitazioni), vi è però una maggiore probabilità di riscontrare problemi d'affidabilità rispetto alla fatica termo meccanica in fase di verifica su prototipo oltre che in esercizio. Non essendo disponibili criteri di progettazione a fatica termo meccanica universalmente accettati, i progettisti industriali tendono a utilizzare approcci basati sull'esperienza per gestire la complessità del problema termo meccanico, confidando su una sperimentazione a banco come fase di verifica imprescindibile del progetto. Di conseguenza è frequente l'allungamento indesiderato della fase di delibera di un nuovo motore con un impatto molto negativo anche dal punto di vista economico.

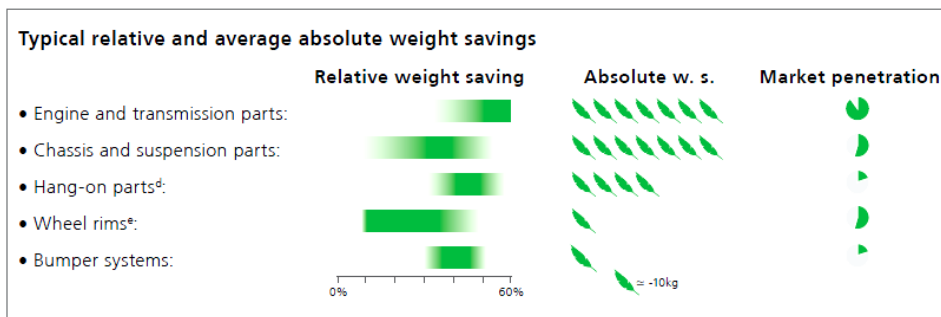


Figura 1. Riduzioni di peso a seguito dell'adozione di leghe d'alluminio per la realizzazione di parti di autoveicolo e livello di diffusione nel mercato. [1]

La crescente competizione nel mercato globalizzato spinge però verso la riduzione del tempo di delibera di un nuovo motore. La riduzione dei tempi di sviluppo complessivo passa attraverso un trasformazione del processo di sviluppo delle singole parti e del motore complessivamente da sequenziale ad integrato secondo le schematizzazioni di Fig. 2

Tradizionalmente le diverse fasi funzionale, termo strutturale e tecnologica della progettazione si sviluppano in modo indipendente e sequenziale. La progettazione del processo realizzativo di una pezzo prende il via dal disegno costruttivo del pezzo finito e si concentra prevalentemente sulla verifica della sua producibilità e qualità macroscopica. Problemi di natura produttiva e/o affidabilistica richiedono quindi correzioni con costose iterazioni. Un approccio integrato caratterizzato dal dialogo tra progettazione e fabbricazione fin da inizio progetto consente di sviluppare da subito e ai vari livelli soluzioni condivise, minimizzando la necessità di iterazioni. Il dialogo è favorito da una elevata integrazione di competenze e di strumenti di progettazione come il CAD, FEA e i simulatori di processo.



Figura 2. a) processo sequenziale di sviluppo di parti fuse
b) processo integrato di sviluppo

Per velocizzare la delibera di un nuovo motore è fondamentale anche ridurre il numero e il tempo dei test a banco. Un contributo significativo è apportato dalla capacità predittiva dal punto di vista dell'affidabilità dei componenti critici, come ad esempio un metodo di calcolo a fatica termo meccanica di teste motore.

L'approccio ideale alla progettazione a fatica termo meccanica si basa sul confronto tra resistenza locale del materiale e le sollecitazioni locali di origine termo meccanica imposte dall'utilizzo potendo individuare così eventuali punti critici in base a criteri di accumulo del danno. Un'implementazione di questo approccio deve però affrontare una serie di ostacoli: i) difficoltà di caratterizzare sperimentalmente il comportamento termo meccanico dei materiali mediante campioni rappresentativi delle proprietà del componente; ii) complessità delle condizioni al contorno per le analisi termo meccaniche e difficoltà di validare l'accuratezza del calcolo; iii) elevata specializzazione delle competenze e difficoltà di interazione tra tecnici analisti e materialisti/tecnologi; iv) difficoltà nel validare sperimentalmente nuove procedure di calcolo.

Questo lavoro sintetizza una nuova procedura di progettazione di una testa motore sviluppata nell'ambito di un progetto finanziato nell'ambito HI-MECH F.A.R tra l'azienda Ferrari di Maranello (MO) leader nella produzione di autovetture sportive e di motori ad elevate prestazioni ed un gruppo di ricerca multidisciplinare. Maggiori dettagli sono riportati in [2-4].

2. PROGETTAZIONE STRUTTURALE DI UNA TESTA MOTORE

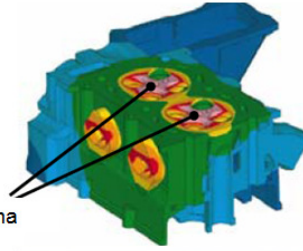
Il progetto ha preso in esame lo sviluppo della testa di un motore 8 cilindri, perché è il componente più critico sia in termini di complessità tecnologico-realizzativa sia come sollecitazioni termo strutturali affaticanti. Le funzioni di una testa motore sono varie e sono evidenti dalla complessità geometrica: garantire la tenuta della camera di combustione, guidare le valvole, immettere ed espellere aria e gas di scarico, posizionare il sistema di iniezione. La complessità della geometria, il numero di parti da produrre e gli obiettivi di leggerezza motivano una produzione per colata in conchiglia. Il materiale tradizionalmente utilizzato è la lega di alluminio da fonderia A356 (G-AlSi7Mg0,3). Il getto è sottoposto al trattamento di tempra di soluzione con invecchiamento artificiale T6 per conseguire la prestazione meccanica ottimale. Le successive lavorazioni meccaniche garantiscono i necessari accoppiamenti e fissaggi degli organi in fase di assemblaggio motore. La testa motore durante il funzionamento è sottoposta a condizioni di sollecitazione complesse: il precarico dato dai tiranti all'atto dell'assemblaggio motore, le dilatazioni differenziali a seguito dei cicli accensione/spengimento, i cicli termici variabili in funzione del tipo di utilizzo, le sollecitazioni dinamiche nella camera di combustione.

La progettazione di una testa motore, oltre alle considerazioni funzionali e tecnologiche, deve garantire l'affidabilità strutturale, cioè un'adeguata resistenza alle sollecitazioni termo meccaniche affaticanti in servizio. Due sono le principali tipologie di criticità termo meccanica riscontrabili in una testa motore mostrate in Fig. 3 assieme ai rispettivi fattori di influenza: i) fatica termica e ii) fatica ad alta temperatura ed alto numero di cicli. Quest'ultima è stata la criticità di particolare interesse per il progetto.

Fatica termica – (basso numero di cicli)

- Alto carico termico localizzato
- Dilatazioni termiche differenziali
- Deformazione plastica localizzata
- Autotensioni termiche al raffreddamento
- Cicli di temperatura (on/off)

Punti critici: piatto fiamma



Fatica ad alto numero di cicli

- Tensione media dovuta alle tensioni di origine termica e al precarico bulloni
- Sollecitazioni oscillanti ad alta frequenza per la pressione di scoppio

Punti critici: intagli geometrici o porosità



Figura 3. Possibili criticità e fattori di influenza legate alla fatica termo meccanica delle teste motore

La nuova metodologia di progettazione integrata di parti motore sviluppata in questo progetto è definita nello schema di Fig. 4. Essa si differenzia dal processo tradizionale per l'introduzione di alcune fasi (i.e. blocchi a contorno tratteggiato) che collegano ed integrano le attività di simulazione di processo con l'attività di calcolo termo strutturale. Pertanto la definizione della geometria della testa basata su considerazioni funzionali dà il via alla progettazione del getto e la simulazione della colata fornisce anche informazioni su microstruttura e difetti attesi nelle diverse parti del getto stesso. L'analisi delle sollecitazioni termo meccaniche col metodo FEM è poi integrata da informazioni sulle proprietà meccaniche locali arrivando ad un calcolo a fatica in tutti gli elementi del modello della testa motore in base ad un criterio predittivo della durata che combina sollecitazioni locali, temperatura e resistenza locale del materiale in presenza di difetti di entità stimata dal processo di simulazione.

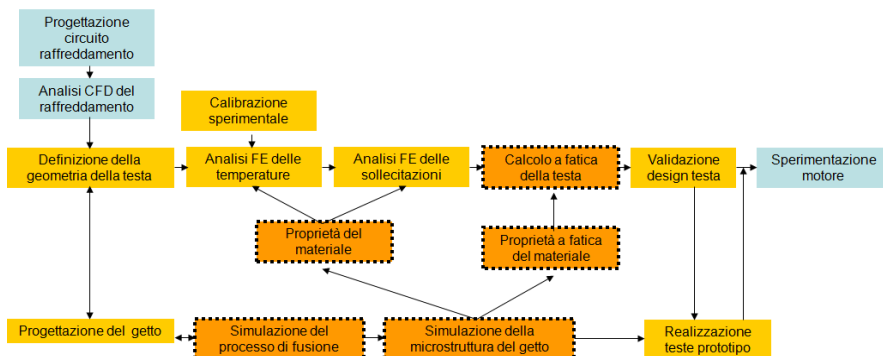


Figura 4. Processo innovativo di sviluppo di getti in lega d'alluminio per applicazioni motoristiche.

Il progetto ha richiesto una partecipazione multidisciplinare per l'interconnessione delle competenze specialistiche che spaziano dalla progettazione /strutturale, metallurgia e tecnologia fusoria ed ha portato allo sviluppo di una serie di attività mirate a: i) la definizione di affidabili criteri di verifica termo meccanica, ii) la disponibilità di una base di dati sperimentali specifici per la lega d'alluminio d'interesse, iii) lo studio dell'interazione tra microstruttura e proprietà meccaniche, con particolare riferimento al ruolo svolto dai difetti di solidificazione nel fenomeno della fatica termo meccanica, iv) la capacità di progettare la colata per simulazione, ottenendo informazioni sulla microstruttura del materiale, tipicamente variabile da zona a zona in un getto di forma complessa. In questo lavoro si delineano i principali aspetti esaminati e si presenta una procedura di validazione metodologica.

3. SOLLECITAZIONI TERMOMECCANICHE DI UNA TESTA MOTORE

Il principale sistema di calcolo delle sollecitazioni termo meccaniche in componenti strutturali complessi è metodo degli elementi finiti (FEM). L'applicazione al caso di interesse inizia con la generazione del modello FEM della testa e degli elementi ad essa collegati. Quindi la metodologia richiede un calcolo preliminare che fornisca il campo termico della testata. Per la determinazione del campo termico cui è soggetta la testa durante il funzionamento si è modellata la testa, la guarnizione, le sedi valvole, le guide valvole, i prigionieri e i dadi di serraggio. Ottenuto il campo termico, è necessario impostare il modello meccanico, immettendo tutti i carichi che i vari componenti scambiano con la testa. L'escursione termica cui la testa va soggetta nel ciclo di accensione-spegnimento è fonte di sollecitazione a bassa frequenza, che può dar luogo a danneggiamento a fatica termica. Nel funzionamento del motore a regime si deve tener conto delle sollecitazioni di tipo termo meccanico dovute ai transitori termici e agli scoppi in camera combustione con cicli di sollecitazione ad frequenza più elevata.

La Fig. 5 evidenzia che la modellazione FEM non ha considerato l'intera testa della testa bensì un modulo delimitato da due piani di simmetria geometrica per ridurre i tempi di calcolo. E' poi mostrato il campo di temperature raggiunte dalla testa a condizione di regime con riferimento alle prove al banco a seguito dell'analisi termica. La zona della testa vicina al piatto fiamma si trova a lavorare a temperature (i.e. 250 °C) elevate per la lega d'alluminio con cui è realizzata, mentre allontanandosi dal piatto fiamma, superando ad esempio il giro acqua superiore, le temperature di lavoro scendono sotto i 150 °C. La modellazione FEM termo strutturale per valutazioni di fatica termica (cfr. Fig. 3) richiede un modello costitutivo del materiale dipendente dalla temperatura. La lega A356 utilizzata allo stato T6 ha caratteristiche meccaniche variabili con la temperatura ed il tempo di esposizione ad alta temperatura, con un comportamento elasto-plastico ben riprodotto dal modello di Chaboche. Il software FEM utilizzato consente di inserire la dipendenza dei parametri del modello in funzione della temperatura. Per simulare direttamente una dipendenza dal tempo e dalla temperatura di invecchiamento cui è soggetto il componente sono state sviluppate specifiche subroutine che correlano tensione di snervamento e durezza Brinell a seguito di specifica sperimentazione, [3]. Si è verificato così che il meccanismo alla base delle rotture che caratterizzano la zona del piatto fiamma delle teste motore è la fatica termica (i.e. elastoplastica a basso numero di cicli).

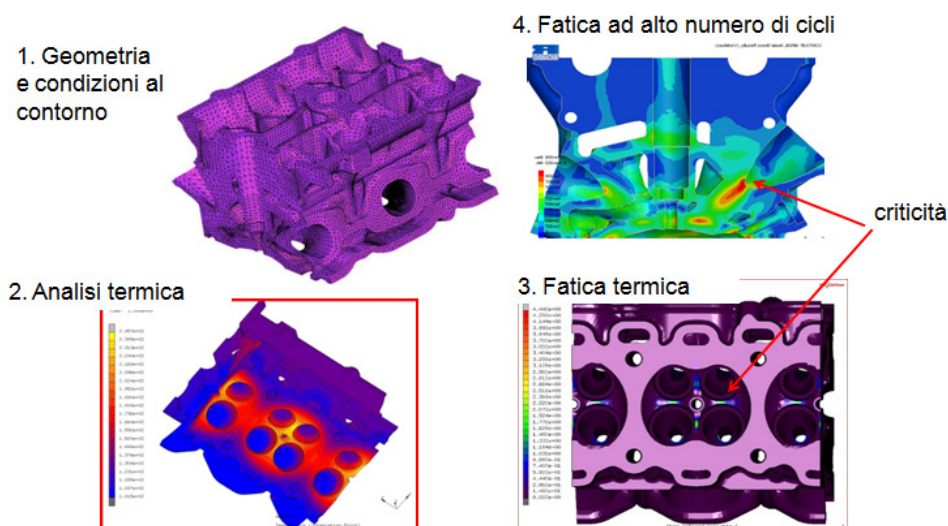


Figura 5. Potenziali criticità identificate con la simulazione FEM termo meccanica, [2]

Nella Fig. 5 si mostra la distribuzione dell'indice di danneggiamento, calcolato dalla subroutine al termine del calcolo. Si nota come esso sia massimo in corrispondenza dei ponticelli tra le valvole, dove si registrano le massime temperature e dove il valore della tensione di snervamento è minima; in queste zone si registrano le plasticizzazioni maggiori. La fatica termica si differenzia da quella classica (in campo elastico ed ad alto numero di cicli) perché utilizza specifici parametri di danneggiamento quale l'energia di deformazione specifica dissipata localmente (i.e. la somma delle aree dei cicli di isteresi delle curve tensione-deformazione). La Fig.5 mostra vi sono altre zone, ad esempio interne alla testa come evidenziato dalla sezione del modello FEM che presentano condizioni di sollecitazione di potenziale criticità per fatica ad alto numero di cicli soprattutto se accoppiate alla presenza di difetti quali pori di dimensioni significative. Nella sezione si evidenzia che il valore della tensione di Von Mises è attorno ai 100 MPa con indice di danneggiamento significativamente elevato.

4. CARATTERIZZAZIONE A FATICA DELLA LEGA PER GETTI

Le proprietà meccaniche di componenti in lega di alluminio da fonderia sono influenzate dalle caratteristiche microstrutturali puntuali e soprattutto dai difetti di solidificazione quali porosità da gas (a forma tendenzialmente sferica), cavità di ritiro (di forma irregolare ramificata) e film di ossido. Nel caso del comportamento a fatica di getti in lega A356 è noto sperimentalmente che sono le porosità ad avere il ruolo più significativo, in quanto discontinuità che fungono da concentratori di tensioni, favorendo la nucleazione e quindi propagazione delle cricche. E' stato inoltre evidenziato come la loro forma, dimensione nonché collocazione all'interno del getto, abbiano influenza, più o meno

marcata, sulla corrispondente vita a fatica, il che conduce ad una notevole dispersione dei dati.

Al fine di quantificare: i) il ruolo dei pori sulla resistenza del materiale, ii) la qualità dei processi di produzione, iii) la capacità predittiva di modelli di fatica basati sulla meccanica dei difetti, si sono condotte molte campagne sperimentali con prove a fatica ad alto numero di cicli (stair-case ridotti con interruzione delle prove a 10^7 cicli). Le prove sono state condotte in buona parte a temperatura ambiente perché nelle teste motore si possono verificare rotture per fatica ad alto numero di cicli in zone soggette a temperatura non molto elevata (i.e. $< 130^\circ\text{C}$). La modalità di sollecitazione sperimentale prescelta è stata la flessione rotante (a 50 Hz). Prove di confronto utilizzando sollecitazioni di trazione-compressione hanno rivelato resistenze inferiori di ca. 15% rispetto al caso della flessione rotante.

I provini di fatica sono stati ottenuti sia per colata diretta in conchiglia di materiali che avevano subito modifiche di lega, sia per estrazione da parti motore ottenute industrialmente e da piastre colate in condizioni controllate, cfr Fig. 6. Come esempio della dispersione della durata a fatica per presenza di porosità in provini di A356 ottenuti per colata in conchiglia in Fig.6 viene presentato un diagramma S/N (ampiezza di tensione vs. numero di cicli a rottura) con i dati di prove flessione rotante e trazione/compressione. Utilizzando l'approccio stair-case ridotto, la resistenza fatica a 10^7 cicli a temperatura ambiente è risultata variare nel campo 40-60 MPa.

Il ruolo centrale delle porosità sulla resistenza a fatica è stato confermato dall'esame delle superfici di frattura dei campioni di fatica. Si è notato così che le rotture hanno avuto sempre inizio da porosità (da gas o cavità di ritiro) come quella di Fig. 7. Sulla superficie si identifica la tipica morfologia di frattura per fatica con innescio in corrispondenza di un poro vicino alla superficie che ha un profilo ovviamente irregolare.

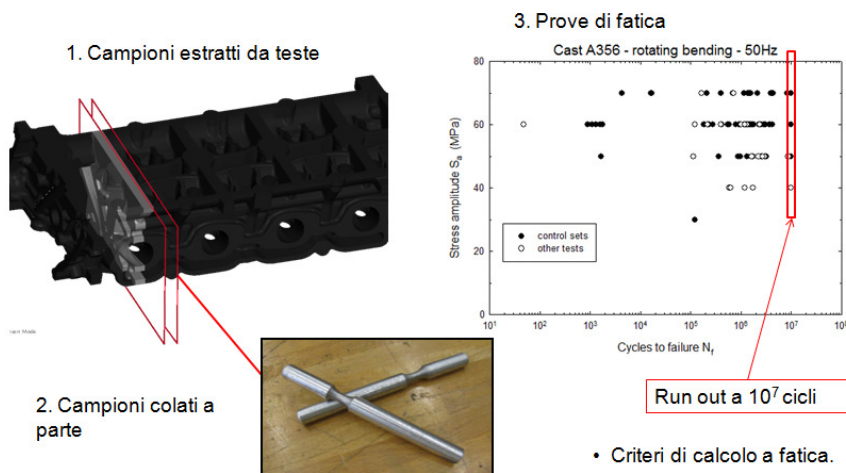


Figura 6. Caratterizzazione a fatica del materiale mediante provini estratti da teste motore o colati a parte

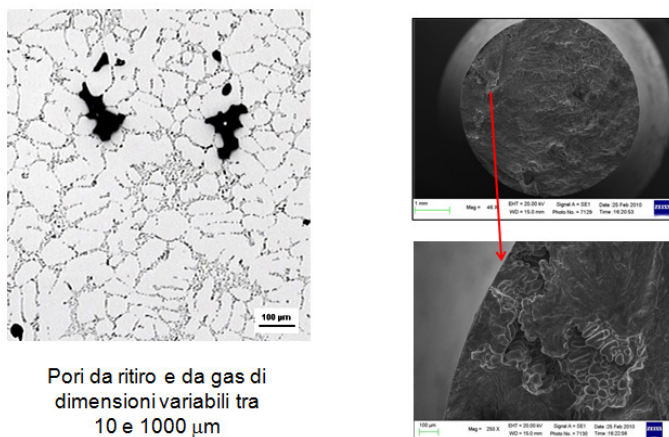


Figura 7. Caratterizzazione a fatica del materiale mediante provini estratti da teste motore o colati a parte

Vi è quindi la zona di propagazione stabile a fatica che si espander dal poro verso l'interno del provino, assumendo progressivamente una configurazione semiellittica, Fig. 7. Raggiunta una profondità critica, dipendente dal livello di sollecitazione e dalla tenacità del materiale, si ha la zona di rottura di schianto finale. I molti dati delle prove di fatica dimostrano chiaramente una forte influenza della dimensione del poro critico, [].

L'evidenza sperimentale è interpretata utilizzando approcci di meccanica della frattura che considerano un poro equivalente ad una fessura e trascurando la fase di innesco della fessura di fatica al bordo poro. In alternativa si può i) stimare la vita a fatica quale numero di cicli richiesto per far propagare la fessura dalla sua dimensione iniziale ad una dimensione finale caratteristica del materiale impiegando leggi di propagazione di tipo Paris e sue varianti specifiche per il materiale ii) stimare la resistenza a fatica in base alla sollecitazione locale calcolata, alla dimensione attesa del difetto critico e al fattore di ΔK di soglia, ΔK_{th} . L'applicazione di questa seconda metodologia al calcolo a fatica di un getto complesso è più agevole e si realizza determinando le sollecitazioni locali al suo interno col FEM, utilizzando un valore di ΔK_{th} rappresentativo della lega utilizzata e ottenendo per simulazione del processo di colata una stima della dimensione massima del poro.

5. SIMULAZIONE DELLA MICROSTRUTTURA DELLA TESTA

Gli obiettivi di questa fase dell'attività sono chiariti dalla Fig. 8 in cui si evidenzia come dalla simulazione del processo di riempimento della cavità con metallo fuso si possano identificare le diverse caratteristiche microstrutturali del getto a seguito delle condizioni locali di raffreddamento e solidificazione.

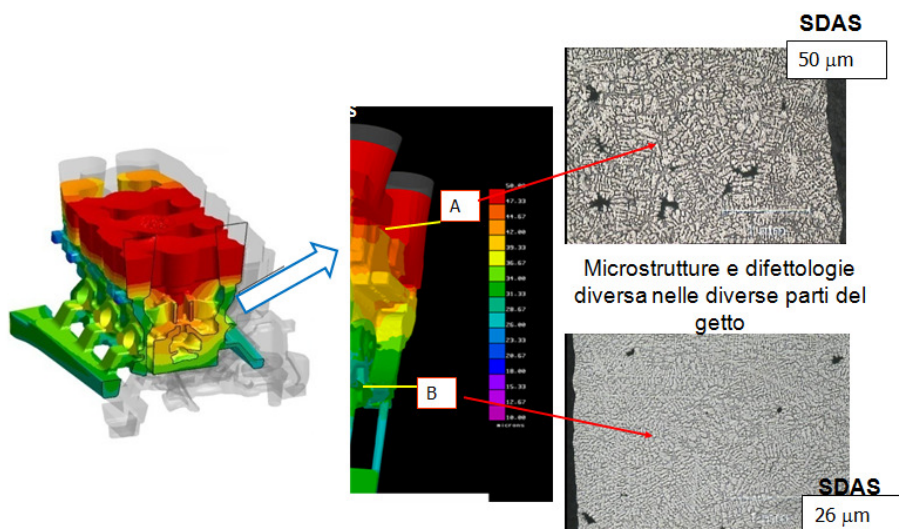


Figura 8. Simulazione del riempimento di uno stampo per colata di lega d'alluminio fusa e mappa dello SDAS. A SDAS minori corrispondono difetti di dimensioni minori.

Alle diverse microstrutture locali si attribuiscono specifiche difettologie, come mostrato qualitativamente in Fig. 8, utilizzando criteri messi a punto durante la sperimentazione su materiali, fatica e difetti descritti nella sezione precedente. Le analisi di riempimento dell'attrezzatura di colata sono state messe a punto e validate mediante osservazione diretta del fenomeno e mediante confronto con dati di processo, con particolare riferimento ai tempi di colata. La successiva analisi del processo di solidificazione è stata portata a termine tramite un fase di aggiustamento dei parametri di simulazione, in particolare dei coefficienti di scambio termico per ottenere una convergenza con i valori di temperatura monitorati in alcuni punti dell'attrezzatura durante il processo.

Le caratteristiche microstrutturali locali sono state quindi determinate sfruttando le note correlazioni tra la velocità locale di raffreddamento e la distanza dei rami interdendritici secondari (SDAS). La messa a punto del sistema previsionale della formazione di porosità nei getti è un tema di grande complessità ed ancora in fase di sviluppo anche per i codici più avanzati. I parametri di nucleazione richiedono una specifica messa a punto per cui si è progettato un getto che consentisse di ottenere una correlazione fra parametri di processo e la porosità. I livelli locali di porosità percentuale nel getto sono stati calcolati mediante un modello previsionale e confrontati sui valori misurati nei getti reali. Da questi getti sperimentali sono stati poi estratti anche campioni per prove di fatica.

Per supportare questo sviluppo metodologico sono state dettagliate analisi microstrutturali su sezioni di una testa motore. Dai campioni metallografici esaminati con un microscopio ottico ed un sistema di analisi d'immagine si sono ottenute mappe di vari parametri quali l'area percentuale dei difetti (Area %), lo SDAS e l'area delle particelle

di silicio eutettico, [3]. L'elevato indice di correlazione fra i valori di SDAS sperimentali e simulati e la coerenza delle curve sperimentali e simulate in tutte le fasi dell'analisi, ha dimostrato l'accuratezza della determinazione dei coefficienti di scambio e l'adeguatezza delle strategie risolutive adottate. Con una mappatura della durezza Brinell e con prove di trazione su provini estratti direttamente da teste si è avuta ulteriore informazione sulle caratteristiche meccaniche locali all'interno di una testa.

6. METODOLOGIA DI VALIDAZIONE

In questa sezione viene presentato l'approccio adottato per validare la metodologia innovativa di progettazione a fatica termo meccanica. Esso fa ricorso all'identificazione di un simulacro rappresentativo della condizioni realizzative della testa motore che sia però più semplice ed economico da produrre e da provare sperimentalmente a fatica.

In Fig. 9 viene mostrata una sezione della testa motore con evidenziato un setto separatore tra canali di raffreddamento con un foro di passaggio. Per effetto delle sollecitazioni termo meccaniche della testa questo setto è potenzialmente critico. Pertanto per il simulacro si è utilizzata la configurazione geometrica locale del setto con l'aggiunta di due estremità piane che consentono un agevole afferraggio da parte di una macchina di prova materiali disponibile in laboratorio. Per la realizzazione del simulacro è stato progettata una attrezzatura per la fonderia che consente la colata simultanea di un elemento da cui si possono estrarre tre parti, cfr. Fig. 9. Il processo di colata del simulacro è stato simulato determinando la mappa delle caratteristiche microstrutturali attese che sono state poi verificate mediante analisi metallografica. La mappa di Fig. 9 evidenzia caratteristiche del materiale, associate ai diversi colori, variabili all'interno di un singolo simulacro e tra un simulacro e l'altro.

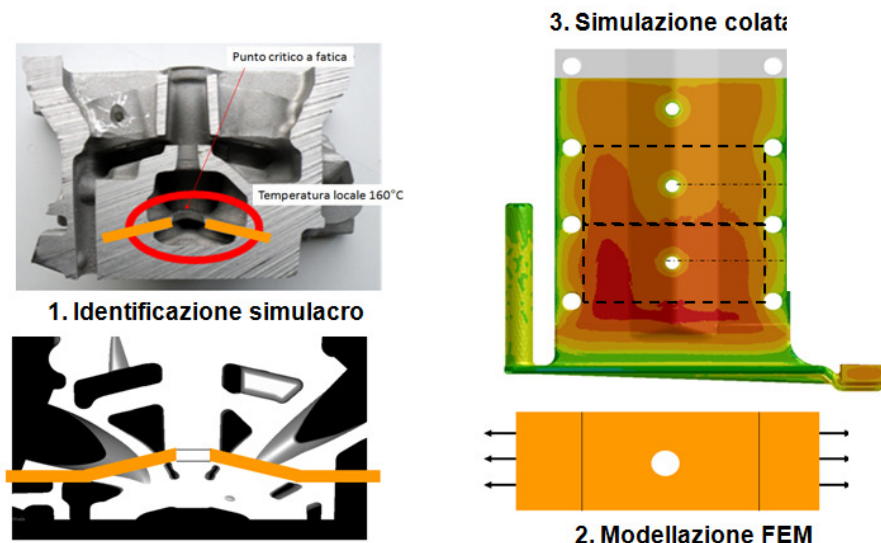


Figura 9. Sviluppo di un simulacro

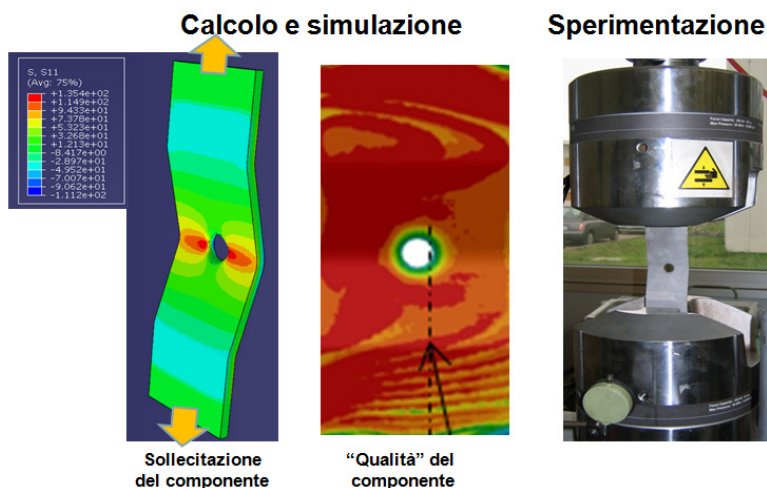


Figura 10. Schema concettuale di validazione della metodologia di calcolo a fatica.

Sulla base delle sollecitazioni ottenute a calcolo FEM che evidenziano specifici punti di concentrazione delle sollecitazioni e della mappa della qualità del materiale (intesa come SDAS e, conseguentemente, la dimensione massima di poro) all'interno del simulacro, cfr. Fig. 10, si è impostata una sperimentazione con test di fatica di ampiezza costante a diversi rapporti di carico che hanno portato alla rottura di vari simulacri. Il confronto tra le condizioni di rottura osservate (i.e. ampiezza della forza applicate e numero di cicli) e quanto prevedibile con il metodo di calcolo messo a punto conferma le potenzialità della nuova metodologia di progettazione integrata termo strutturale.

7. CONCLUSIONI

Si è presentata un'evoluta metodologia di calcolo a fatica termo meccanica di teste motore. La metodologia si caratterizza per l'integrazione dei risultati della simulazione del processo fusorio con quelli della simulazione termo meccanica per migliorare ed accelerare il processo di sviluppo. Una prima introduzione della metodologia in ambito industriale è avvenuta grazie ad un progetto con collaborazione multisettoriale.

RINGRAZIAMENTI

Si ringraziano doverosamente l'ing. Gianluca Pivetti di Ferrari ed i colleghi Lorella Ceschini, Antonio Strozzi e Luca Tomesani ed i rispettivi collaboratori nelle diverse sedi universitarie per aver partecipato al progetto FAR qui sintetizzato.

BIBLIOGRAFIA

[1] Anon. 2007, Aluminium in cars, European Aluminium Association

- [2] Nicoletto G. et al. 2010, “Progettazione integrata di componenti motore in lega di alluminio ottenuti per fusione e sollecitati a fatica ad alta temperatura”, Procs. I° Convegno Nazionale del Coordinamento della Meccanica Italiana, Palermo, Paper No. 47, 12 pp.
- [3] Ceschini L. et al. 2010, “Influenza della microstruttura di solidificazione sulla resistenza a fatica di leghe Al-Si-Mg da fonderia”, Procs. I° Convegno Nazionale del Coordinamento della Meccanica Italiana, Palermo, Paper No. 46, 10 pp
- [4] Tomesani L. et al. 2010, “Modellazione di processo e previsione della microstruttura nella colata in conchiglia di testate motore in alluminio”, Procs. I° Convegno Nazionale del Coordinamento della Meccanica Italiana, Palermo, Paper No. 44, 12 pp

NORMALIZATION OF THE STRESS CONCENTRATIONS AT THE INTERFERENCE FIT BETWEEN A CYLINDRICAL SHAFT AND A HUB WITH ROUNDED EDGES

Antonio Strozzi

*Department of Mechanical and Civil Engineering,
University of Modena and Reggio Emilia, Italy
E-mail: antonio.strozzi@unimore.it*

Luca Bertocchi

*Department of Mechanical and Civil Engineering,
University of Modena and Reggio Emilia, Italy
E-mail: luca.bertocchi@unimore.it*

Eleonora Campioni

*Department of Mechanical and Civil Engineering,
University of Modena and Reggio Emilia, Italy
E-mail: eleonora.campioni@unimore.it*

Abstract. *This paper concerns the stress field at the extremities of a press-fit connection between a shaft and a hub, in which the indentation is mitigated by the presence of a filleted edge.*

Although the classical interference-fit Lamé solution accurately predicts the stress field as soon as the end-effect perturbation decays - and holds in average, more detailed analyses are required in order to evaluate the stress concentrations at the rounded edge.

Derived from analytical considerations, a normalizing parameter is proposed that accounts for the combined effect of various geometric and material parameters, thus making the compilation of FEM derived design charts viable.

The configuration in which an infinitely long solid shaft is press-fitted within a hollow hub is considered in detail, and a stress concentration diagram for various inner-to-outer radii ratios is provided. This diagram supports the designer in the selection of the correct fillet radius.

Keywords: *contact mechanics, interference-fit, stress concentration factor*

1. INTRODUCTION

Interference fits are widely employed to connect gears, pulleys, flanges, wheels, and similar mechanical components, to a shaft. When a cylindrical shaft of infinite length is press-fitted into a cylindrical hub of finite length, stress concentrations take place at the hub-shaft contact extremities, whereas the contact stresses remain reasonably constant along the central portion of the contact, e.g. references [1-5]. If the edges of the hub bore are sharp,

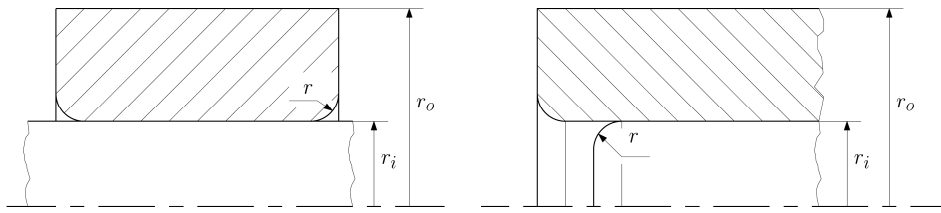


Figure 1. Infinitely long shaft (a), and partially inserted shaft (b).

the elastic pressure peaks are mathematically infinite, whereas they become finite when the corners are rounded. To introduce fillets at the hub bore edges is therefore a practically relevant means for reducing the hub stress peaks.

Additional methods adopted to relieve the pressure peaks consist in introducing a shoulder in the shaft, and in adopting a grooved hub or shaft, see e.g. references [1,6,7].

While the stress state at the hub sizeable central zone may be confidently evaluated by modelling the press-fit problem as plane and by employing the Lamé equations for thick-walled cylinders, e.g. references [1,7], the localised pressure peaks are not amenable to a simple analytical evaluation, e.g. references [2,5]. Consequently, the transmittable torque may be reliably estimated by resorting to the Lamé-based plane solution, that is valid along the majority of the contact axial length; conversely, a detailed, specific three-dimensional analysis is needed to evaluate the local stress peaks and to assess the strength of the hub.

In the following, it is assumed that the rounded edges of the hub are described by a quarter of circumference, so that the fillet radius fully describes the geometry of the mating surfaces at the hub-shaft contact extremities.

Fig. (1) depicts the reference assembly, showing the two allowed configurations of a shaft-hub press-fit: in (1a) an infinite shaft is completely inserted within a finite length hub, whereas in (1b) configuration a semi-infinite shaft is partially inserted within a semi-infinite or infinite hub.

It is also noted that the contact pressure peaks depend on the shaft radius r_i , the hub outer radius r_o , the fillet radius r , the diametrical interference I , and the hub and shaft Young's moduli, which are supposed equal for the sake of simplicity. Instead, the hub axial length is generally hardly relevant, provided that the hub is sufficiently long to produce, in it, central part, a uniform stress state in the axial direction; this situation implies that the two pressure peaks at the contact extremities do not interact among each other.

The normalised contact pressure peaks would depend on the previous radii normalised over a reference radius, but even so the number of variables seems to be too high to allow design charts to be compiled, that can cover a wide range of hub-shaft geometries and of materials adopted. In this paper a viable way to compile design charts addressing the stress peaks is presented for an infinitely long, solid cylindrical shaft, press-fitted into a cylindrical hub of finite axial thickness with rounded edges.

In particular, it is shown in this study that, for a prescribed ratio between the hub inner and outer radii, the normalised stress, peaks in the presence of radiussed edges of the hub bore depend upon a parameter that comprises and summarises the simultaneous effects of a reference radius for the shaft-hub press-fit, the fillet radius, the interference, and the Young's modulus. For a prescribed hub radial aspect ratio, it is therefore now possible to prepare design charts in which the normalised von Mises stress peaks are diagrammatically

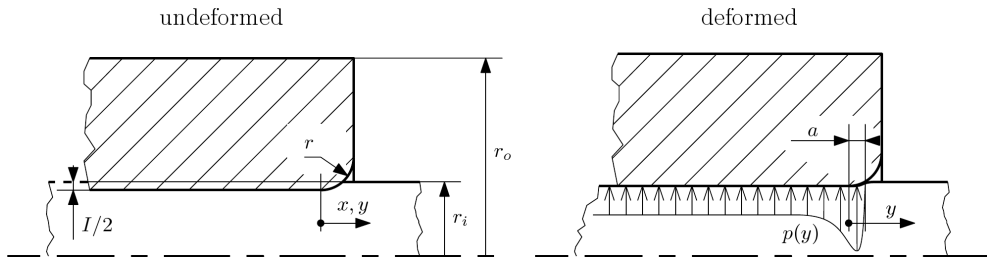


Figure 2. Original and deformed shape.

presented versus the above normalising parameter.

2. NORMALIZING PARAMETER RATIONALE

It is assumed in this study that the hub-shaft contact is frictionless, since the coefficient of friction is low in this kind of applications, and the direction of the shear stresses acting at the hub-shaft contact in the axial direction is often undefined, since it depends on the assembly procedure. Consequently, fretting fatigue aspects, e.g. reference [8], are not examined in this study. The materials of the hub and shaft are assumed to exhibit the same Young's modulus. Finally, the contact problem is assumed as elastic, and yielding of the materials is not considered.

In this paper the above normalising parameter is determined by following the same approach successfully adopted in references [9,10], referring to a pin-in-plate contact problem, and to a press-fit problem, respectively.

To normalise the integral equation, the initial clearance/overlapping must be approximated, as closely as possible, by a function formed by the product of a shape function by a representative variable. This is a key point of this approach, see e.g. references [9,11].

To this aim, the contact problem is formally expressed in terms of the integral equation

$$\int_{-\infty}^{+a} k(x, y) p(y) dy = \frac{I}{2} - r + \sqrt{r^2 - x^2} \quad (1)$$

where the total elastic radial deformation due to shaft contraction and hub expansion is set equal to the interference pointwise value, within the contact area.

The known term of the integral equation represents the initial clearance and the overlapping between the two undeformed mating profiles of the hub and shaft.

It appears convenient to simplify the right-hand side of Eqn. (1) by employing a parabolic approximation for the circular arc, and hence obtaining

$$\int_{-\infty}^{+a} k(x, y) p(y) dy = \frac{I}{2} - \frac{x^2}{2r} \quad (2)$$

Provided that the edge rounding is considered as a contact profile correction only, and its influence on the hub stiffness is neglected, the problem is cylindrical in nature and the Green function k is expected to depend only on the radial aspect ratio r_i/r_o .

Therefore we have

$$k(x, y) = \frac{\bar{k}}{E} \left(\frac{x}{r_i}, \frac{y}{r_i} \right) \quad (3)$$

where the kernel is normalised with respect to the Young's modulus and a radius.

A further step in normalizing requires the definition of a reference interference pressure, which can be taken as the unperturbed value far from the terminals; such p_0 value can be expressed according to [16] as

$$p_0 = \frac{EI}{2r_i} \frac{1}{1 + \frac{r_i^2 + r_o^2}{r_i^2 - r_o^2}} = \frac{EI}{r_i} f\left(\frac{r_i}{r_o}\right) \quad (4)$$

Finally, the normalized form for (2) can be written as

$$\int_{-\infty}^{+\frac{a}{r_i}} \bar{k}\left(\frac{x}{r_i}, \frac{y}{r_i}\right) \frac{p\left(\frac{y}{r_i}\right)}{p_0} d\left(\frac{y}{r_i}\right) = \frac{1}{2f\left(\frac{r_i}{r_o}\right)} \left\{ 1 - \left(\frac{x}{r_i}\right)^2 \left[\frac{r_i^2}{rI} \right] \right\} \quad (5)$$

from which it appears evident that if two press-fit problems possess the same r_i/r_o radii aspect ratio and the same adimensional ratio

$$\phi = \frac{rI}{r_i^2} \quad (6)$$

they share the same normalized solution for the contact problem.

The normalizing parameter ϕ summarizes the simultaneous effect of the fillet radius, r , of the shaft radius, r_i , and of the interference I . The Young modulus appears in the p_0 definition only.

The formulation of the title problem in terms of an integral equation serves to define the normalizing parameter ϕ ; however, since the integral equation is quite complex, e.g. references [2,5], stress field is more conveniently evaluated by the finite element method, e.g. reference [12], by varying the normalising parameter within an interval covering practically significant shaft-hub press-fit configurations, e.g. reference [13]. The introduction of a normalising parameter allows the stress forecasts to be presented in a way that is both more compact and of ample validity, with respect to papers analysing specific problems, e.g. references [14,15].

3. NUMERICAL CALCULATIONS FOR DESIGN CURVE RETRIEVAL

Calculations have been performed with the commercial FE package MSC.Marc/Mentat 2010; an axisymmetric model has been used in order to concentrate the computational effort toward finer meshes, since very high local stress gradients are predicted in the vicinity of the rounded edge. Linear triangular elements are used for a better compliance with the

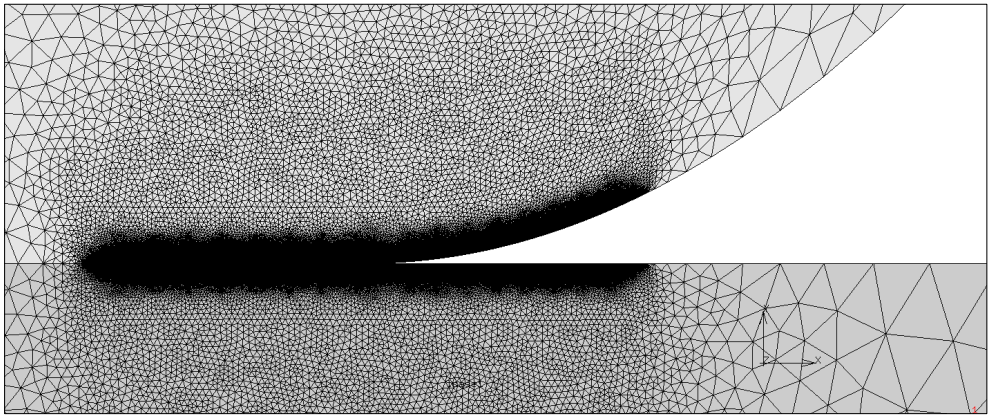


Fig. 3. Detail view of the FE model discretization.

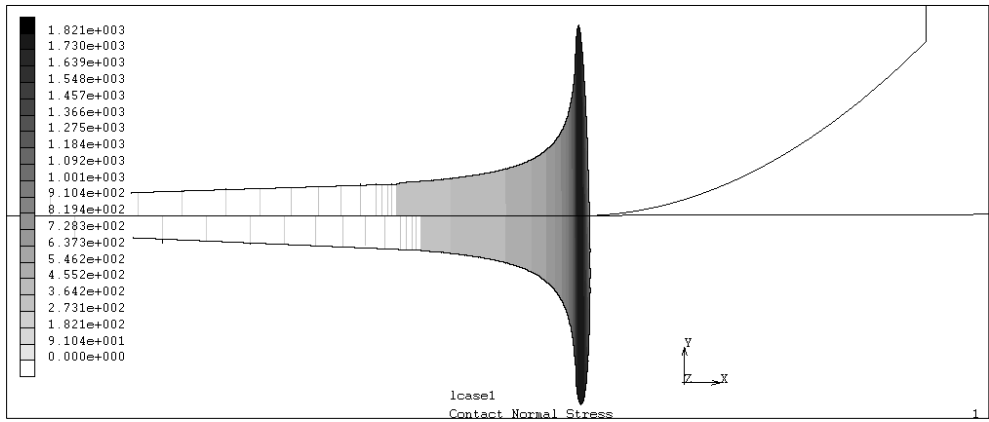


Figure 4. Pressure profile for the test case (scale in MPa).

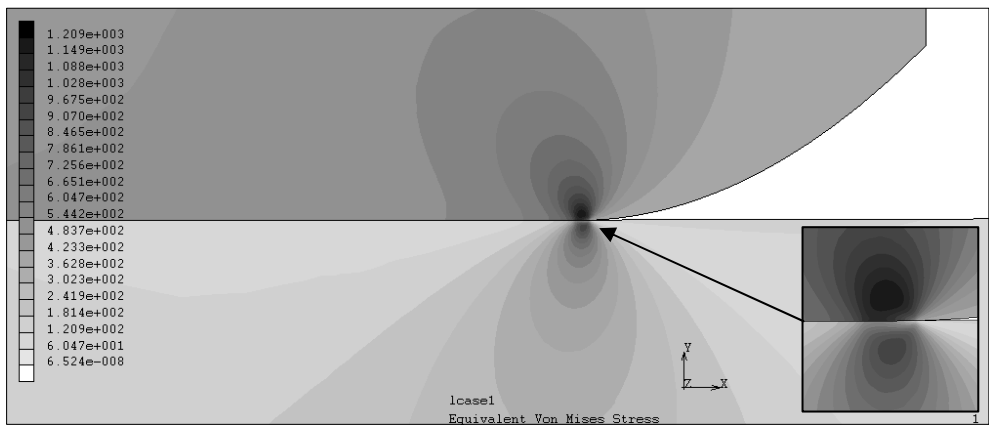


Figure 5. Equivalent von Mises stress map for the test case (scale in MPa)

contact algorithm, and an element edge length varying from $0.00025 \cdot r$ to $0.001 \cdot r$ along the area of interest. A sample discretization is shown in Fig. (3).

Being the problem non-linear due to the contact area variation, the coupling is modelled as neat-fit at the initial condition, whilst a simulated radial thermal expansion of the shaft is used to obtain a growing interference value. The axial component of a fictitiously orthotropic thermal expansion relation is profitably exploited in order to reduce the discretization misalignment at the reciprocal surfaces, which is due to the Poisson effect.

An example FE output result (test case: $r_i=20\text{mm}$, $r=1\text{mm}$, $l=0.1\text{mm}$, $r_i/r_o=0.5$) is presented in Fig. (4) in terms of contact pressure profile; the related von Mises equivalent stress is shown in Fig. (5). It can be noted that in similarity with the Hertz theory the equivalent stress reaches a maximum at a subsurface point.

4. DESIGN CURVES

In Fig. (6), a preliminary design diagram is presented that express the normalised stress concentration factor versus the above normalising parameter for a selection of hub radial aspect ratios; such concentration factor is defined as the maximum von Mises in the hub divided by the reference pressure p_0 .

The curves appears well approximated by a power law regression formula

$$K_t = a\phi^{-b} \quad (7)$$

where a and b are parameters which depend on inner-to-outer radii ratios r_i/r_o , as in Tab. (1). Such diagram assist the designer in selecting the correct radius for the rounded edge of the hub, or in assessing the hub stress level in the shaft-hub press-fit.

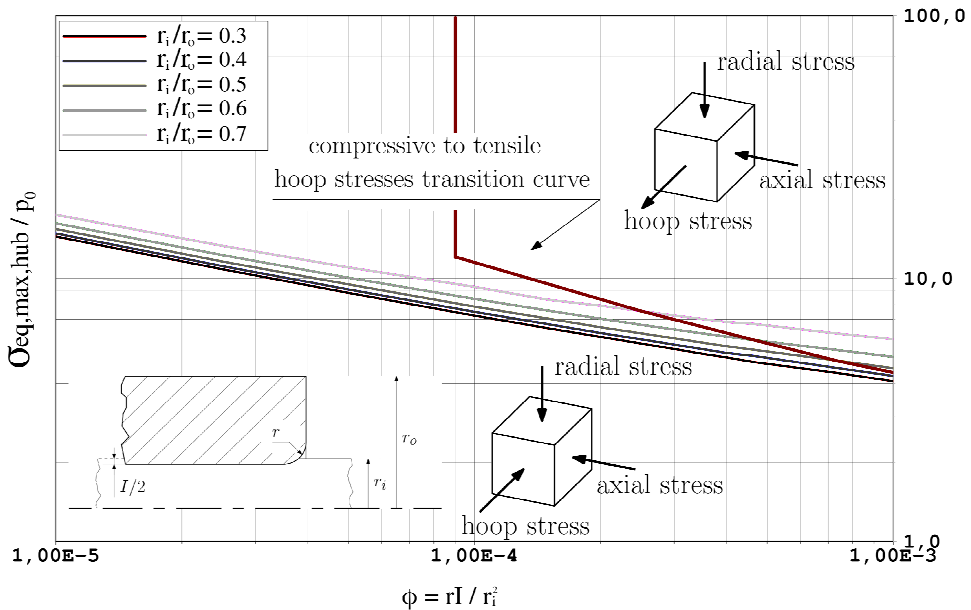


Figure 6. Bilogarithmic diagram of K_t for various r_i/r_o ratii.

Table 1. Numerical parameters for Kt interpolation curves

r_i/r_o	a	b	error, %
0.3	0.567	0.279	-43 ÷ +24
0.4	0.614	0.274	-44 ÷ +24
0.5	0.680	0.2686	-44 ÷ +35
0.6	0.804	0.258	-51 ÷ +33
0.7	1.041	0.242	-59 ÷ +53

5. NOTES ON THE HOOP STRESS NATURE IN THE HUB

The hoop stress in the hub, sampled at the point in which the equivalent stress is maximum, switches from tensile to compressive depending on the ϕ parameter. The classical interference fit theory predicts tensile hoop stresses, which are observed in the higher ϕ range, i.e. large interference values and large fillet radii and therefore larger contact areas along the rounding. However, in the lower ϕ range, compressive hoop stresses have been obtained beneath the pressure peak, where the von Mises stress is maximum.

Such unexpected behaviour can be rationalized considering that for the smaller fillet radii the singular pressure condition is approached, in which the radial stress is minus infinite; the hoop stress has to approach minus infinite too in order to maintain the circumferential strain finite, according to

$$\varepsilon_{\theta} = \frac{u}{r} = \frac{1}{E}(\sigma_{\theta} - \nu(\sigma_r + \sigma_a)) \quad (8)$$

and being the axial stress negligible.

Moreover, as depicted in Fig. (7), the material beneath the pressure peak is constrained in expansion by the underlying understressed zones, and approaches an in-plane hydrostatic stress condition similar to a rubber ring within an axially unrestricted housing.

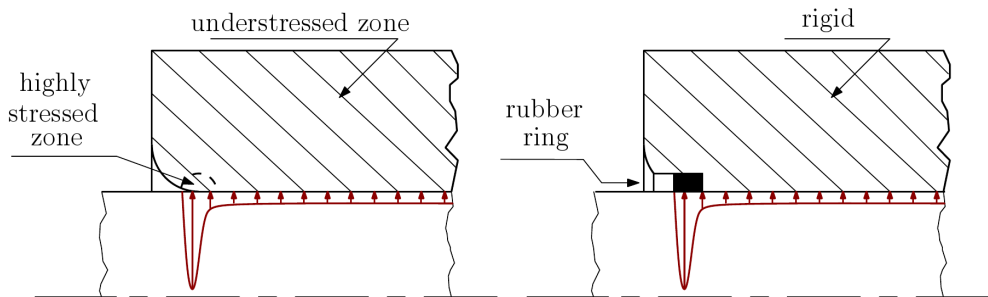


Figure 7. Rationalization of the compressive hub stress condition in low ϕ models

6. CONCLUSIONS

A normalizing parameter ϕ has been defined that summarizes the simultaneous effects of the fillet radius, shaft radius, interference, and Young's modulus.

The parameter ϕ allows compact diagrams to be prepared that assist the designer in evaluating the stress concentrations in a shaft-hub press-fit.

REFERENCES

- [1] Kanber, B., 2006. "Boundary element analysis of interference fits". *Turkish Journal of Engineering and Environmental Sciences*, **30**, pp. 323-330.
- [2] Goodier, J.N., Loutzenheiser, C.B., 1965. "Pressure peaks at the ends of plane strain rigid die contacts (elastic)". *Journal of Applied Mechanics*, **32**, pp. 462-463.
- [3] Conway, H.D., Farnham, K.A. , 1967. "Contact stresses between cylindrical shafts and sleeves". *International Journal of Engineering Science*, **5**, pp. 541-554.
- [4] Strozzi, A., 1986. "Static stresses in an unpressurized, rounded, rectangular, elastomeric seal". *ASLE Transactions*, **29**, pp. 558-564
- [5] Ciavarella, M, Hills, D.A., Monno, G. , 1998. "The influence of rounded edges on indentation by a flat punch". *Proceedings of the Institution of Mechanical Engineers: Part C*, **212**, pp. 319-328.
- [6] Bijak-Zochowski, M., Marek, P., Tracz., M., 1994. "On methods of reduction and elimination of stress singularities in some elastic contact problems". *International Journal of Mechanical Sciences*, **36**(4), pp. 279-296.
- [7] Strozzi, A., *Costruzione di Macchine*, 1998, Pitagora, Bologna.
- [8] Lanoue,, F., Vadean, A., Sanschagrin, B., 2009. "Finite element analysis and contact modelling considerations of interference fits for fretting fatigue strength calculations". *Simulation modelling, practice and theory*, **17**, pp. 1587-1602.
- [9] Ciavarella, M., Baldini, A., Barber, JR., and Strozzi, A. , 2006, "Reduced dependence on loading parameters in almost conforming contacts". *International Journal of Mechanical Sciences*, **48**, pp. 917-925.
- [10] Strozzi, A., Baldini, A., Giacomini, M., Rosi, R., Bertocchi, E. , 2009. "Contact stresses within a split ring inserted into a circular housing". *Journal of Strain Analysis*, **44**, pp. 2041-3130.
- [11] Schmelz, F., Seher-Thoss, H.C., Aucktor, E. , 1992. *Universal Joints and Driveshafts*. Springer-Verlag, Berlin.
- [12] White, D.J., Humpherson, J., 1969. "Finite-element analysis of stresses in shafts due to interference-fit hubs". *Journal of Strain Analysis*, **4**, pp. 105-114.
- [13] Castagnetti, D., Dragoni, E., 2005. "Optimal aspect ratio of interference fits for maximum load transfer capacity". *Journal of Strain Analysis*, **40**, pp. 177-184.
- [14] Hang, Y., McClain, B., Fang, X.D., 2000. "Design of interference fits via finite element method". *International Journal of Mechanical Sciences*, **42**, pp. 1835-1850.

- [15] Dobromirski, J., Smith, I.O. , 1986. "A stress analysis of a shaft with a press-fitted hub subjected to cyclic axial loading". *International Journal of Mechanical Sciences*, **28**, pp. 41-52.
- [16] Croccolo, D., Vincenzi, N., 2009. "A generalized theory for shaft--hub couplings", *Proceedings of the Institution of Mechanical Engineers: Part C*, **223**, pp. 2231-2239.

SPERIMENTARE PER PROGETTARE

Alessandro Freddi

Dipartimento di Ingegneria Meccanica e Aeronautica,

Università di Bologna, Italia

E-mail:alessandro.freddi@unibo.it

Sommario. *Il progetto, attività centrale e distintiva dell'ingegnere, richiede lo sviluppo di strumenti di carattere logico-matematico e sperimentale, necessari per affrontarne gli aspetti pragmatici, sintattici e semantici. In particolare sono essenziali strumenti per la modellazione matematica dei dati e strumenti logici per la chiarificazione dell'informazione che è di natura empirica, ottenuta direttamente dall'utente o indirettamente dal mercato. In questo modo l'invenzione e la scelta delle soluzioni progettuali sono correttamente guidate. Nel presente lavoro si pongono in luce in particolare, le fasi che si avvalgono della sperimentazione per interpretare i requisiti e il contesto, cioè gli aspetti pragmatici e per ottimizzare le scelte strutturali, cioè gli aspetti sintattici.*

Parole chiave: *la progettazione ingegneristica, gli strumenti logico-matematici, il metodo sperimentale.*

1. INTRODUZIONE

Il lavoro seguente prende avvio da tre considerazioni:

- La progettazione è strumento d'innovazione,
- La progettazione è attività caratterizzante dell'ingegneria,
- La progettazione si avvale del metodo sperimentale.

Tali affermazioni sembrano ovvie ma in realtà la scienza dell'ingegneria si è mossa prevalentemente in altre direzioni e non sempre si considera l'attività e l'apprendimento della progettazione come distintiva rispetto alle scienze pure. Oggi si assiste a un interesse accademico prevalente sul versante delle scienze analitiche sia nella ricerca che nella formazione, rispetto agli aspetti sintetici tipici della progettazione.

Non solo: troppo spesso il progetto è legittimato solo quando è espresso in termini di analisi, come si rileva dalle espressioni che rientrano nei libri di progettazione:

- Analisi dei materiali
- Analisi di fatica
- Analisi di meccanica della frattura
- Analisi di affidabilità
- ecc.

Queste analisi sono molto importanti ma non mettono in luce e danno per scontati i contenuti compositivi/sintetici della progettazione.

Inoltre il progetto è visto a volte come processo di sviluppo di un sistema mentre altre volte identifica il risultato di questo processo che si realizza in un artefatto, spesso frutto di un notevole lavoro intellettuale e pratico, raramente considerato di valore scientifico perché non è formalizzato in modo matematico. L'attività di progettazione è quindi vista come poco rigorosa e frutto più di un'esperienza come una espressione artistica, che di una ricerca scientifica, [1]. Con questo pregiudizio, imparare a progettare si può solo acquisire

progettando, cioè con un'attività pratica. Questo crea anche un divario crescente tra gli interessi dell'Accademia e quelli industriali.

In realtà il progetto è un processo che richiede operazioni combinate di analisi e di sintesi, come recita il testo seguente, che definisce il progetto secondo l'Ente di Accreditamento degli ingegneri negli Stati Uniti, [2]:

“ Design is the process of devising a system, component, or process to meet desired needs. It is a decision making process (often iterative), in which the basic sciences, mathematics, and engineering sciences are applied to convert resources optimally to meet a stated objective. Among the fundamental elements of the design process are: the establishment of objectives and criteria, synthesis, analysis, construction, testing and evaluation.....”

La grande difficoltà del progetto risiede nel fatto che esso possiede più dimensioni che si estendono a diverse competenze e specializzazioni, individuate, già nell'antichità, nelle tre categorie: “Utilitas, Firmitas, Venustas” [3], le quali possono pensarsi attinenti rispettivamente alla pragmatica, alla sintassi e alla semantica, cioè alla funzione, alla struttura e alla forma del prodotto.

Il presente lavoro si propone di rilevare un contributo di scientificità anche in quelle fasi del processo di progettazione che vengono generalmente considerate frutto dell'esperienza e quindi di scarso valore scientifico. Si vuole mostrare come, almeno per gli aspetti funzionali e strutturali, si debba applicare il metodo sperimentale che oltre ad essere il fondamento delle scienze analitiche, assolve a questo compito anche per gli aspetti sintetici della progettazione. Per gli aspetti formali ed estetici occorre infine ricordare che i canoni dell'estetica sono di natura descrittiva e non prescrittiva e quindi possono solo essere acquisiti con un'educazione critica che faccia riferimento a principi generali [22].

Per il metodo sperimentale vale quanto espresso sinteticamente nell'affermazione seguente, [4]: *“Fin dalla nascita della scienza sia l'ipotesi come la verifica sperimentale dell'ipotesi hanno carattere matematico. [Infatti] sia nell'ipotesi sia nell'esperimento la realtà è espressa matematicamente, in quanto ridotta ad un insieme di grandezze fisiche misurabili (non qualitative)...”*

Questa osservazione evidenzia il contenuto matematico della sperimentazione che, come verrà dimostrato, non va confusa con una mera anche se accurata, collezione di dati empirici, cioè di “prove”.

2. IL METODO DI PROGETTAZIONE

I metodi di progettazione trovano il loro fondamento nel metodo cartesiano. E' interessante notare che a questa conclusione, gli studiosi sono pervenuti in modo indipendente, sia per la progettazione ingegneristica, [5], sia per il “Design” industriale [6].

Schematizzando, il processo di progettazione si può immaginare articolato in fasi che rispecchiano i punti salienti del “Discorso sul metodo” cartesiano [7]. La sequenza sintetica, (circolare come proiezione di una struttura elicoidale del processo che riproduce l'aspetto iterativo della progettazione), è descritta in Fig. (1) nei punti seguenti:

Definizione del compito:

- Individuare opportunità e pianificare il prodotto
- Chiarificare il compito e verificare la fattibilità di nuove tecnologie
- Specificare i requisiti tecnici

Sviluppo del progetto concettuale:

- Concepire nuove soluzioni

Sviluppo del progetto costruttivo:

- Dimensionare, cioè dare corpo alle soluzioni concettuali

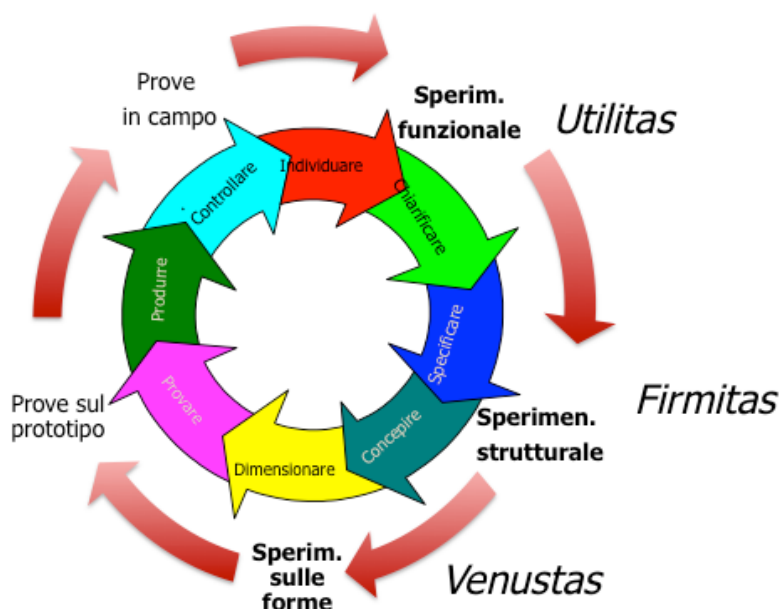


Figura 1. Uno schema a otto fasi rappresentativo del processo di progettazione.

- Eseguire prove sul prototipo e completare il progetto di dettaglio
- Sviluppo del progetto produttivo:
- Organizzare e sviluppare la produzione
 - Controllare il prodotto nel suo ciclo di vita.

Linguaggio comune e linguaggio tecnico

Una volta pianificato il prodotto, si enumerano i requisiti che esso deve avere. Le esigenze e i requisiti del potenziale utilizzatore sono espressi nel linguaggio comune e quindi sono in genere aggregati e interrelati, cioè confusi. Il primo compito è quindi di chiarificazione, per tradurli in parametri tecnici misurabili cioè in variabili matematiche.

Le interrelazioni dei requisiti sono spesso rappresentate con diagrammi a blocchi puramente descrittivi che non evidenziano il livello di dipendenza tra i requisiti stessi, né il livello d'importanza relativa. In termini cartesiani si può affermare che essi non costituiscono un repertorio d'idee chiare e distinte, perché essi esprimono spesso concetti confusi e ambigui, Fig. (2).

Nell'esempio di Fig. (2) è riportata l'analisi delle esigenze espresse da potenziali utenti per la definizione delle caratteristiche di un motociclo. Esse sono espresse in modo confuso e interrelato cioè non sono "chiare e distinte".

Secondo un autorevole parere, l'acquisizione e l'uso di strumenti di chiarificazione sarà proprio uno dei compiti principali dell'ingegnere del futuro che non dovrà essere preparato solo al trattamento di problemi algoritmici, già chiari nelle variabili in gioco, nei quali la soluzione certamente esiste anche se non nota [8].

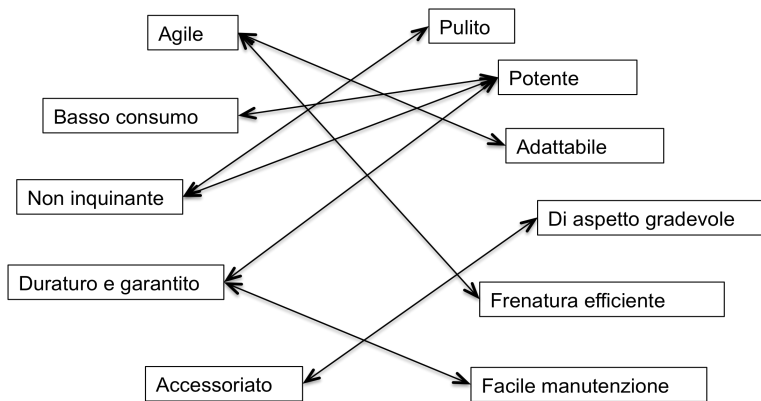


Figura 2. Le esigenze come emergono da analisi di mercato per il progetto di un motociclo.

3. COME SPERIMENTARE LA "UTILITAS"

Le funzioni di "utilitas" di un sistema cioè i requisiti formalizzati matematicamente, una volta chiariti, vengono trattati con i noti metodi matematici di tipo analitico. Nel problema di natura meccanica essi riguardano generalmente gli aspetti cinematici e dinamici.

Esiste però un problema che in parte precede l'analisi funzionale del sistema, presente però anche lungo fasi successive, che richiede il trattamento e la classificazione di dati empirici con modelli logico-matematici e quindi impiega strumenti che si inquadrano pienamente per quanto detto, nel metodo sperimentale.

Tali strumenti servono a risolvere problemi di interpretazione, quali:

- Valutazione logica delle dipendenze
- Valutazione delle criticità dei punti più deboli
- Valutazione statistica delle interrelazioni.

Essi servono rispettivamente nella fase di sviluppo del progetto e di guida nelle scelte di tutti i parametri che lo costituiscono, precisamente:

- Per il trattamento e la chiarificazione del flusso d'informazioni
- Per l'eliminazione o, almeno, la riduzione delle criticità e dei malfunzionamenti
- Per lo sviluppo di progetti di esperimenti sul sistema.

Valutazione logica delle interrelazioni tra i dati empirici

Si fa riferimento alle matrici logiche d'interrelazione e di relazione che servono a tradurre i requisiti dell'utente nei requisiti tecnici ovvero nei parametri di specifica tecnica.

Per esempio, con matrici d'interrelazione si mettono in relazione i requisiti (le esigenze espresse e non espresse) dell'utilizzatore con loro stessi, disponendoli secondo le righe e secondo le colonne e interpretando i requisiti in colonna come cause e gli stessi di riga come effetti. La dipendenza fra i requisiti può essere nulla, debole, media, forte e può essere quantificata, ad esempio, rispettivamente con valori convenzionali 0, 1, 3, 9.

L'aspetto interessante del metodo consiste nel fatto che esso fa emergere i pochi, in generale, requisiti di colonna che hanno influenza predominante e, dalle somme per colonna, quelli più importanti, (ossia i reali requisiti che hanno influenza prevalente sugli altri) e, per riga, i requisiti che, viceversa, hanno una forte dipendenza dai primi, Fig. (3).

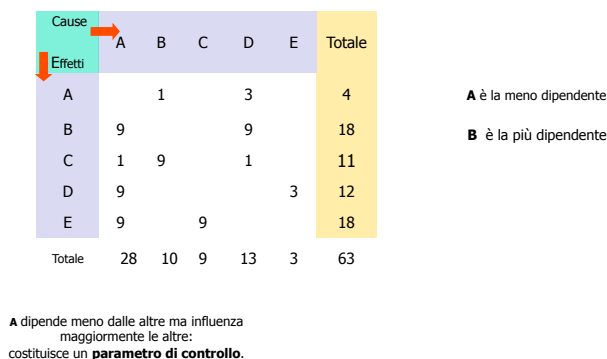


Figura 3. La matrice d'interrelazione.

Viceversa le matrici di relazione pongono in relazione i requisiti più importanti così selezionati, con i parametri tecnici che appartengono viceversa al patrimonio conoscitivo dei tecnici. Esse differiscono da quelle d'interrelazione perché con le matrici di relazione vengono poste a confronto le esigenze ricavate da analisi del mercato (o dal potenziale utilizzatore del prodotto), ordinate secondo i criteri di massima influenza o di massima dipendenza rispetto alle altre, con i parametri tecnici che soddisfano quelle esigenze assieme ai loro valori numerici. Le matrici d'interrelazione sono quindi principalmente uno strumento di marketing mentre le matrici di relazione sono strumenti di "ingegneria".

Il metodo che sintetizza i risultati di queste elaborazioni, noto con l'acronimo QFD (Quality Function Deployment), è in grado di chiarire tutti gli aspetti di un progetto ottimale. Il termine "qualità" non deve trarre in inganno: non rappresenta un valore aggiunto perché ogni processo di progettazione deve porsi la meta di un risultato di qualità. Lo strumento si articola nella composizione di matrici d'interrelazione e di relazione che fissano i "paletti" di una corretta progettazione che consentono di tenere conto di ciascun elemento importante.

Lo strumento QFD fu proposto con l'ambizione di condurre in modo molto formalizzato tutto il processo di progettazione; questa estensione non corrisponde all'esperienza e al pensiero dell'autore, specialmente nella progettazione normale che riguarda lo sviluppo di soluzioni adattative o di variante e molto raramente (forse di un caso su cento) di soluzioni totalmente originali. La forma che assume il complesso delle matrici viene denominata col termine: "*Casa della Qualità*" Fig. (4), con un manifesto riferimento all'origine culturale giapponese del diagramma e del metodo. Il quadro chiarisce le informazioni su:

- esigenze dell'utente classificate in ordine di importanza (righe)
- confronto tra le esigenze e i parametri tecnici necessari per soddisfarle
- confronto con la concorrenza dal punto di vista dell'utente
- confronto con la concorrenza dal punto di vista delle soluzioni tecniche
- direzioni delle scelte per innovare le soluzioni.

E' quindi un metodo ideale per preparare la fase di concezione vera e propria del progetto. Questo è il senso di uno strumento che a buon motivo può essere classificato come

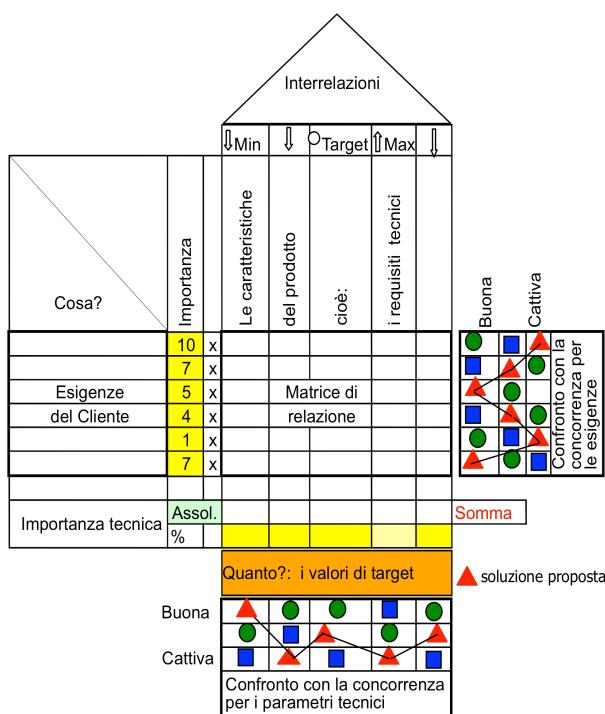


Figura 4. La casa della qualità.

strumento di sintesi, basato sulla logica e su ampie analisi di dati empirici, [9].

Il metodo ha anche un valore aggiunto perché s'inquadra nella filosofia organizzativa del progetto indicata con il nome d'ingegneria concorrente. Essa si propone l'applicazione di un metodo sistematico di progettazione di prodotto che prenda in considerazione tutti gli elementi del ciclo di vita, dalla fase di concezione a quella di uso, con definizione simultanea del prodotto, dei processi di fabbricazione e di tutte le fasi connesse con il suo ciclo di vita. In coerenza con il metodo QFD, l'ingegneria concorrente mette in conto tutte le caratteristiche desiderate nella fase iniziale del processo, per produrre un progetto più "robusto", cioè in grado di sopportare le variazioni del processo produttivo e dell'uso, a costo inferiore del processo sequenziale.

Valutazione delle modalità di danno

Sempre nella stessa prospettiva, lo strumento che raccoglie, in uno schema logico, i dati sperimentali ricavati dall'analisi del nuovo progetto come dall'esperienza su progetti simili, è la FMEA (Failure Mode and Effect Analysis)

La FMEA è una *tecnica di analisi preventiva* utilizzabile per il progetto di un prodotto o di un processo, al fine di assicurarsi che, per quanto possibile, tutti i punti deboli, cioè tutte le potenziali modalità di guasto siano prese in considerazione e sia valutato il rischio relativo, sulla base della gravità delle conseguenze, della frequenza delle cause e

dell'efficacia delle azioni di controllo e siano poste in essere misure per contenerle in limiti accettabili.

Si può considerare come un'applicazione marginale del metodo sperimentale in quanto elabora dati empirici per costruire modelli logici, in modo concettualmente simile al metodo della QFD. A differenza dell'albero dei guasti, che è uno strumento solo logico, la FMEA (o FMECA come viene indicata in una variante tale strumento di controllo del rischio), richiede il trattamento anche di dati empirici, legati all'esperienza dei tecnici e può essere impiegata nella fase concettuale e nella fase costruttiva, [9].

Il Progetto dell'esperimento (DOE) e le superfici di risposta

Questo strumento indicato, in lingua inglese, con il termine "Design of Experiment" (DOE), è basato sulla pianificazione statistica di esperimenti mirati e offre l'inquadramento ideale di ogni tipo di sperimentazione che intenda fare chiarezza su un processo e su un prodotto.

In particolare nelle fasi di progettazione concettuale e costruttiva, il metodo è utile per la ricerca dei valori ottimali delle variabili che influenzano il comportamento del sistema, cioè dei parametri di controllo, in quanto con questo metodo si identificano quali di tali parametri, detti "fattori" e quali loro combinazioni, ottimizzano il risultato del processo in termini di affidabilità, di riduzione dei costi o di qualsiasi altro "indice di qualità".

La Fig. (5) mostra lo schema di un processo con variabili di controllo all'ingresso, variabili in uscita (indici di qualità) e variabili di disturbo che rappresentano l'influenza dell'ambiente assieme alle variabili non prese in considerazione. Nella Fig. (6) a sinistra è mostrato un esempio di risultato dell'esperimento. Sul piano orizzontale è riportato il dominio di definizione dei fattori di controllo in forma normalizzata tra i valori -1, +1. Le quattro ordinate corrispondenti sono i valori sperimentali della risposta del sistema. Quando le variabili sono continue, anche la risposta varia in modo continuo e si rappresenta con una superficie che interpola i quattro valori. Essa è detta superficie di risposta. Naturalmente per un numero di fattori superiore a due è una superficie a più dimensioni che non può essere rappresentata.

L'esempio di sinistra mostra una superficie di risposta con effetto d'interazione, cioè di non linearità, (la superficie che interpola i punti sperimentali si distacca da quella piana che rappresenta un comportamento lineare tra i fattori). Sulla destra è mostrato un esempio di superficie di risposta con forti effetti d'interazione, [10]. In questo caso i punti sperimentati, cioè i valori dei fattori, sono i quattro corrispondenti agli spigoli del dominio più i quattro punti a metà dei lati più il punto centrale del dominio.

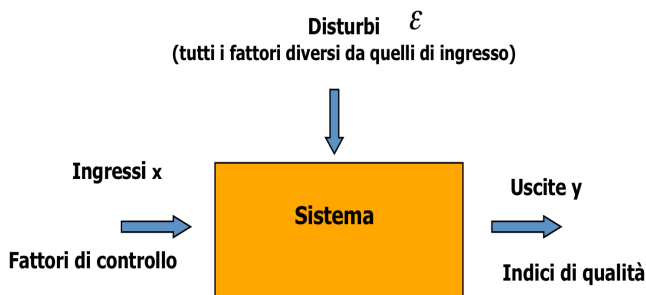


Figura 5. Lo schema del progetto dell'esperimento

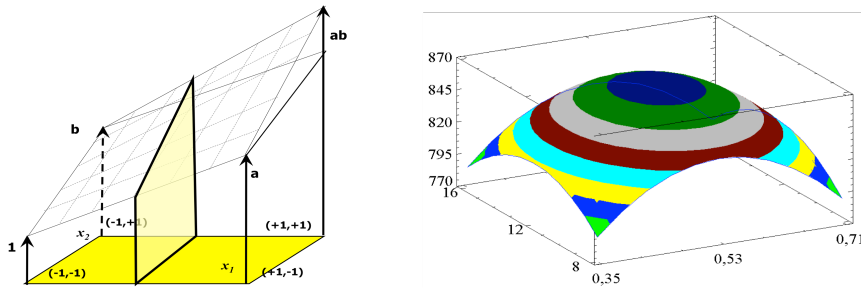


Figura 6. Le superfici matematiche interpretano i rilievi sperimentali ma offrono la risposta del sistema su tutto il dominio delle variabili.

4. COME SPERIMENTARE LA “FIRMITAS” DELLE STRUTTURE

Con riferimento al progetto strutturale, la sperimentazione sulla resistenza rappresenta il campo più noto di applicazione del metodo sperimentale al progetto. Come si è detto nell'introduzione, esso è uno dei settori d'interesse prevalente della ricerca e della formazione.

Anche se più noti, vale tuttavia la pena discutere la particolare tipologia dei problemi strutturali che si avvalgono di dati sperimentali perché pongono il concetto molto importante di problema mal posto. Il fondamento dell'analisi strutturale è riposto nella nota formulazione della legge di Hooke: "Ut Tensio, sic Uis" (come l'estensione, così la forza) che, a una variabile o a più variabili, come oggi viene estensivamente applicata con i metodi numerici agli elementi finiti, si presenta secondo il formalismo delle (1) per il problema diretto e per il problema inverso e lo schema matriciale di Fig. (7).

$$y = A \cdot x \quad e \quad x = A^{-1} \cdot y \quad (1)$$

Il problema, risolto in modo diretto procedendo dalle forze x note alle “estensioni” (deformazioni) y incognite, presume la conoscenza del sistema A e si avvale dei metodi numerici di risoluzione, idonei a strutture complesse.

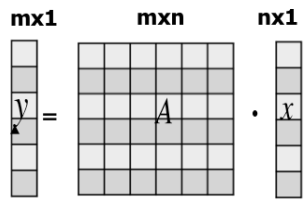


Figura 7. La legge di Hooke per problemi strutturali lineari elastici a una o più variabili. Ad una variabile x e y sono numeri, a più variabili x e y sono vettori.

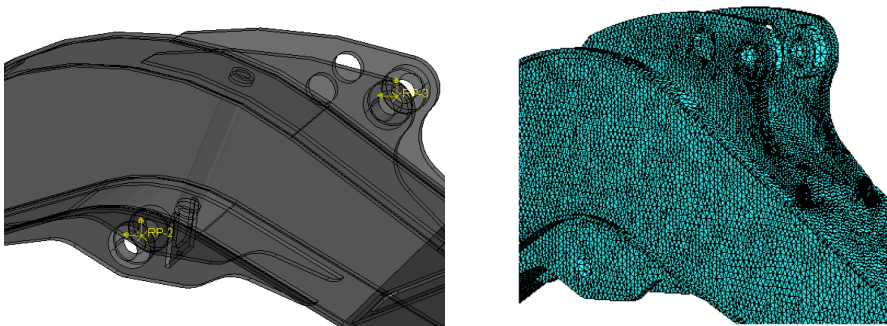


Figura 8. Gli elementi finiti riducono i corpi di forma complessa a un insieme di elementi semplici, per i quali si può scrivere una relazione matriciale del tipo indicato nella Fig. (7).

L'esempio della Fig. (8) illustra una estensione della legge di Hooke; un caso con migliaia di variabili [11] per la descrizione del comportamento strutturale di un braccio di gru. Raramente però il problema si presenta in questi termini nei casi reali: in genere o le forze in gioco o i parametri del sistema non sono noti e la soluzione procede dalla misura delle deformazioni alle variabili incognite. E' quindi un problema inverso.

E' importante osservare che la sperimentazione genera sempre un problema inverso [12] [13], che richiede un modello matematico per interpretare i dati di misura e quindi determinare le incognite, Fig. (9).

I problemi inversi possono essere di due tipi:

- dati gli effetti y e date le cause x si cercano i parametri del sistema A (incognite) oppure:
- dati gli effetti y e dati i parametri del sistema A si cercano le cause x (incognite)

Esistono tre tipologie di problemi inversi, a seconda che il numero degli effetti sia uguale, superiore o inferiore al numero delle cause. I problemi risultano rispettivamente determinati sovra e sotto determinati, con metodi di soluzione appropriati.

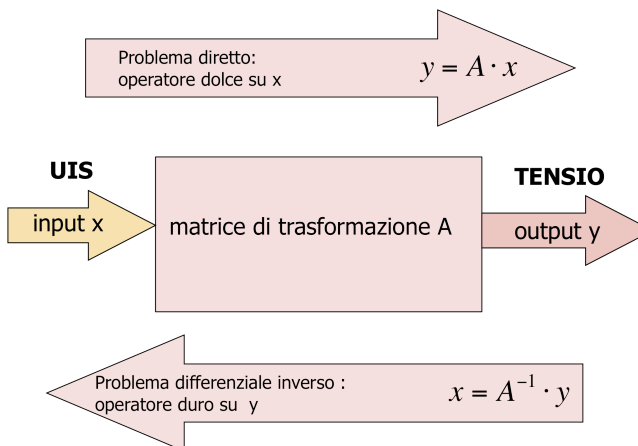


Figura 9. Problemi diretti e problemi inversi, [14].

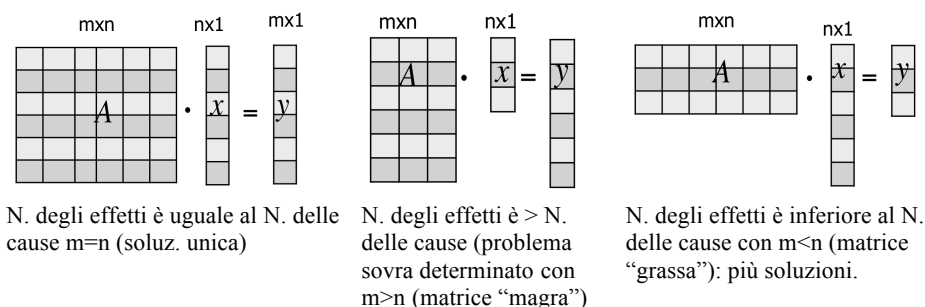


Figura 10. Cause (UIS) ed effetti (TENSIO).

Ricordando il significato di problema mal condizionato, cioè quando piccole variazioni sui dati danno grandi variazioni sulle stime delle incognite, è bene ricordare dalla trattazione teorica che sia i problemi diretti sia gli inversi possono essere mal condizionati.

In particolare per i problemi inversi accade quanto descritto nella Fig. (10):

- un problema determinato può essere o meno mal condizionato
- un problema sovra determinato può essere o meno mal condizionato
- un problema sotto determinato è sempre mal condizionato.

I problemi sovradimensionati e sottodimensionati possono essere risolti in modo approssimato. Infatti la matrice del sistema non è quadrata e non possiede in termini rigorosi una matrice inversa. Può essere invertita in modo approssimato. Come si vedrà dagli esempi, è necessario introdurre il concetto di matrice pseudo inversa.

Esempi di applicazione di problemi strutturali inversi

Si faccia riferimento al più comune tra i metodi di analisi sperimentale dei materiali e delle strutture, basati su analogie fisiche mostrati sinteticamente nella Fig. (11), che è il metodo estensimetrico.

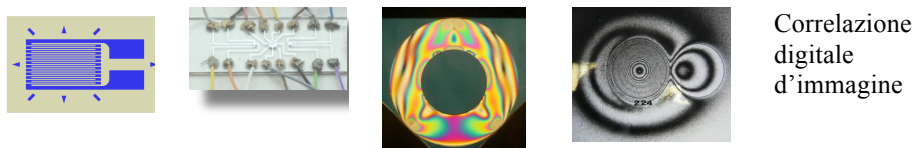


Figura 11. Metodi sperimentali basati sulla misura dell'energia elettrica al posto dell'energia meccanica e sulla misura delle proprietà ottiche al posto di quelle meccaniche. Le prime due immagini a sinistra rappresentano due tipi diversi di estensimetri elettrici, resistivi e piezoresistivi.

Progetto di un telaio motociclistico [15]. Sia considerato noto il sistema con un primo dimensionamento del telaio e siano incognite le forze applicate. Il progetto è articolato nelle seguenti fasi:

- La taratura di un elemento (ammortizzatore) per la misura delle forze trasmesse mediante l'applicazione di deformazioni imposte, (primo problema inverso).
- La misura con prove “in campo” delle reali forze in gioco, per un protocollo di missione stabilito, (secondo problema inverso).
- La determinazione delle reali deformazioni del telaio dovute a queste forze, (problema diretto).
- L'ottimizzazione della forma con la ripetizione del calcolo diretto su più casi (non presentata in questo articolo), per portare il valore delle deformazioni in tutte le parti del telaio a valori corretti.

Il primo punto richiede la “regolarizzazione” dei dati sperimentali ottenuti da misure con estensimetri elettrici sullo stelo dell'ammortizzatore, che collega al telaio la forcella posteriore Fig. (12) al centro, condotte in laboratorio con una macchina universale di prova. L'andamento regolare e crescente della nube dei dati sperimentali Fig. (12) a destra, (il problema è sovra determinato), suggerisce l'impiego di un modello lineare tra forze e deformazioni nello stelo che, con il metodo dei minimi quadrati viene adattato ai dati. Nella relazione lineare seguente sono i parametri a e b di A le incognite approssimate:

$$y = A \cdot x$$

$$J(a,b) = \|y^{meas} - A \cdot x\|^2 = \sum_i \left[\frac{y_i^{meas} - a - bx_i}{\sigma_i} \right]^2 \quad \begin{aligned} \frac{\partial J(a,b)}{\partial a} &= 0 \\ \frac{\partial J(a,b)}{\partial b} &= 0 \end{aligned} \quad (2)$$

Il secondo punto è illustrato nella Fig. (13). In questo caso il problema inverso è determinato e l'equazione seguente dà la relazione tra le deformazioni misurate nello stelo e le forze in gioco, essendo già stati determinati i parametri lineari del sistema “ammortizzatore”.

$$x^{stim} = A^{-g} \cdot y^{mis} \quad (3)$$

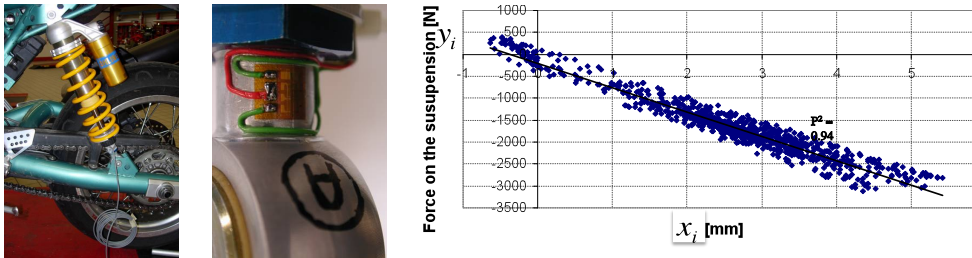


Figura 12. La calibrazione di un elemento di misura delle forze (dinamometro) mediante la conoscenza di deformazioni imposte. Al centro lo stelo strumentato, a sinistra il rilievo.

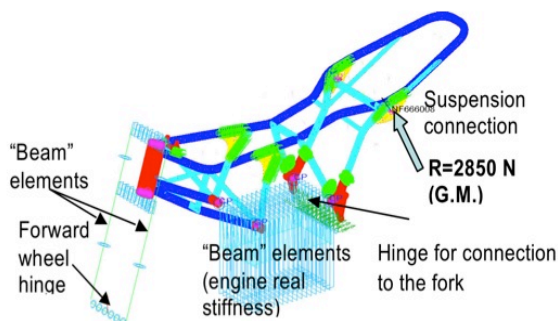
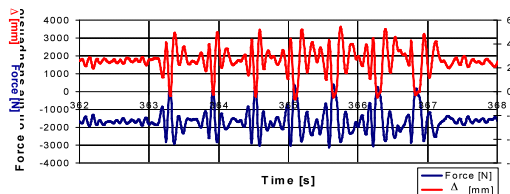


Figura 13. Confronto tra deformazioni nello stelo dell'ammortizzatore e forze trasmesse dalla forcella (cioè dalla ruota), al telaio e, in basso, andamento delle deformazioni nel telaio dovute alle forze così determinate.

In questo caso la matrice è una matrice inversa esatta, non pseudo inversa. Nella Fig. (13) in alto sono posti a confronto gli andamenti delle deformazioni misurate e delle forze corrispondenti, ottenute per il percorso su una pista con ostacoli definiti.

Le deformazioni nel telaio sono determinate infine dalle forze ricavate in precedenza e dalla conoscenza dei parametri del sistema "telaio" fissati mediante un dimensionamento di tentativo, con la soluzione di un problema diretto. L'immagine della Fig. (13) in basso mostra il quadro riassuntivo dell'andamento delle deformazioni nella forma di aree di livello. Come è stato anticipato, la forma ottimale del telaio è stata ottenuta con una serie di calcoli diretti [15], con un andamento simmetrico delle deformazioni nei due lati del telaio, distribuzione considerata ottimale, stante la presenza di un solo ammortizzatore.

Problema inverso determinato ben condizionato [16]. La costruzione di una cella di carico a sei gradi di libertà presentata in Fig. (14) è un esempio di applicazione di un problema sperimentale inverso per la determinazione dei sei parametri di forza e di momento trasmessi dallo sci allo scarpone. La prima soluzione è ottenuta con un numero d'incognite pari al numero di misure, come si evince nello schema a destra che riporta, per chiarezza, solo i quattro anteriori dei sei trasduttori di forza. Il problema è quindi determinato ben condizionato e non ha generato nessun problema d'interpretazione dei dati.

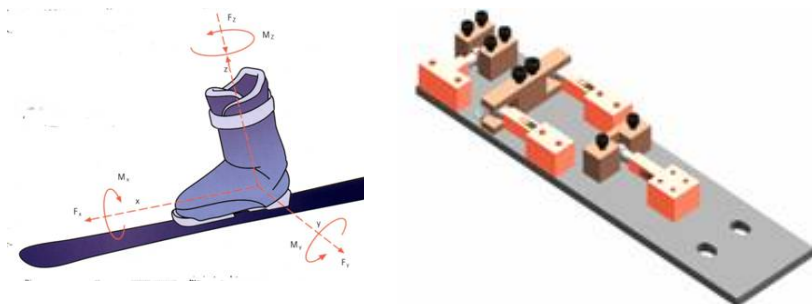


Figura 14. Cella di carico a sei gradi di libertà con sei incognite e sei misure. Problema inverso determinato ben condizionato.

Problema inverso sovradeterminato con il metodo dei minimi quadrati. Viceversa una seconda soluzione, nata per motivi di sicurezza con una struttura iperstatica con sei incognite e otto misure sperimentali, è riportata nella Fig. (15).

Il modello matematico seguito è il seguente; data l'espressione generale dei minimi quadrati:

$$J = \|y^{meas} - A \cdot x\|^2 = \min \quad (4)$$

dalla condizione di minimo:

$$\frac{\partial J}{\partial x} = 0 \quad -A^T (y^{meas} - Ax) = 0 \quad (5)$$

si ricava il sistema di equazioni normali seguenti:

$$A^T y^{meas} = A^T A x^{estim} \quad (6)$$

dalle quali si ottiene l'espressione della matrice pseudoinversa mal condizionata (7):

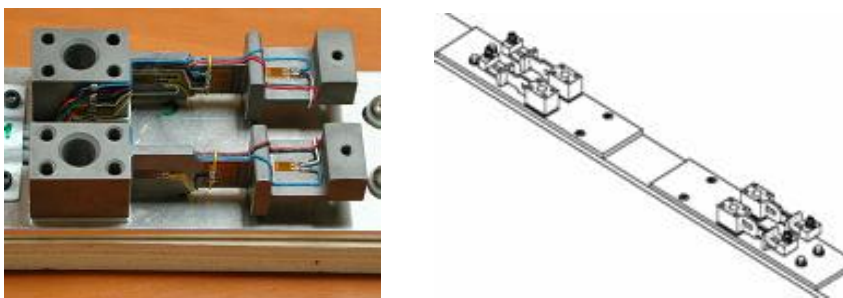


Figura 15. Cella di carico a sei gradi di libertà con otto misure sperimentali

$$A^{-g} = (A^T A)^{-1} \cdot A^T$$

$\begin{matrix} nx1 \\ x^{stim} \end{matrix}$

$=$

$\begin{matrix} nxm \\ A \end{matrix}$

\cdot

$\begin{matrix} mx1 \\ y^{meas} \end{matrix}$

(m > n)

$$x^{stim} = A^{-g} \cdot y^{mis}$$
(7)

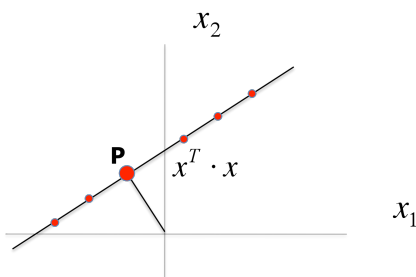
In questo caso la soluzione adottata ha creato alcune difficoltà connesse con il mal condizionamento della matrice del sistema e ha finito per essere scartata nell'applicazione pratica a favore della prima che ha rivelato un eccellente comportamento durante tutte le prove.

Problemi inversi sotto determinati (m < n) (sempre mal condizionati)

Il problema indeterminato è più raro nella meccanica strutturale classica. Esso deriva p. e. da problemi di manutenzione e di analisi di degrado di grandi manufatti nei quali il numero dei dati sperimentali misurabili difficilmente può pareggiare o superare il numero delle incognite, costituite in questo caso dai numerosi parametri del sistema o dalle forze in gioco, [12]. Il problema ammette infinite soluzioni. Per renderlo determinato esistono più modi.

Soluzione di lunghezza minima. La funzione da minimizzare è un funzionale composto da due parti, legate da un moltiplicatore di Lagrange: una prima che seleziona tra le infinite soluzioni quella a distanza minima dall'origine e una seconda che rappresenta la condizione di minima distanza tra i dati misurati e il modello teorico. Dalla condizione di minimo del funzionale globale si ricava l'espressione (8) della pseudo matrice inversa.

$$J = x^T \cdot x + \lambda^T \cdot (y^{meas} - Ax) = \min \quad A^{-g} = A^T (A \cdot A^T)^{-1} \quad (8)$$



$$a_1 x_1 + a_2 x_2 = y^{meas}$$

$$x^{stim} = A^{-g} \cdot y^{mis} \quad (9)$$

Figura 16. Matrice pseudo inversa per un problema inverso sotto determinato. Condizione di lunghezza minima.

Metodo di regolarizzazione di Tikhonov-Tjhung-Miller. Un altro metodo, utile anche per problemi determinati ma mal posti, consiste nel comporre, con la condizione di minimo classica, una condizione di regolarizzazione ottenuta con l'introduzione di una matrice di numeri che è basata su proprietà di tipo teorico della soluzione attesa. Oltre alla matrice di regolarizzazione si aggiunge un coefficiente numerico o una seconda matrice di numeri α che dà ulteriori gradi di libertà nella scelta della soluzione ottimale, [12], [13].

Un esempio di problema determinato ma mal posto che richiede questo tipo di regolarizzazione è il seguente, nel quale le misure di deformazione con tre estensimetri elettrici, intendono rilevare le tensioni residue agenti in un corpo che abbia per es. subito un trattamento termico, quando queste tensioni interne vengono rilassate con un'operazione meccanica come quella di foratura, [17].

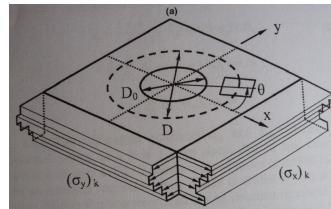
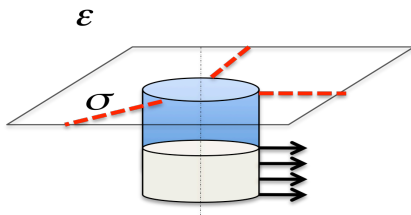
Il problema è complesso e sicuramente mal condizionato, perché si vuole determinare la distribuzione di tensione lungo tutto lo spessore di foratura ma l'effetto di questa tensione "liberata" sui sensori disposti sulla superficie strumentata, diventa sempre più debole all'approfondire del foro. L'espressione della funzione da minimizzare è la seguente nella quale sono indicate le deformazioni e le tensioni con simbologia evidente:

$$J = \left\| \varepsilon^{meas} - A \cdot \sigma \right\|^2 + \alpha \left\| R \cdot \sigma \right\|^2 = \min \quad (10)$$

Per definire R occorrono considerazioni di natura teorica sulla soluzione, [12]: se p.e., la variazione locale di σ (tensione residua) può essere approssimata con una retta, si ha un minimo sulla derivata seconda, il che porta ad una espressione di R come nella (12) seguente:

$$\sigma(\xi) \text{ linear in } \xi: \min \left(\frac{\partial^2 \sigma}{\partial \xi^2} \right) \quad \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix} \quad (12)$$

La matrice R è tridiagonale con il numero delle righe pari al numero dei passi di foratura e il/i fattore/i α controllano l'entità della regolarizzazione. Nella Fig. (17), oltre allo schema della misura, sono riportate le espressioni della relazione che lega le misure alle incognite e l'espressione della matrice pseudo inversa ottenuta con la minimizzazione del funzionale J.



$$\sigma^{estim} = A^{-g} \cdot \varepsilon^{meas}$$

$$A^{-g} = (A^T A + \alpha \cdot R^T R)^{-1} \cdot A^T \quad (11)$$

Figura 17. Metodo ASTM: Misura delle tensioni residue lungo lo spessore, misurate mediante rilassamento ottenuto con un foro cieco, [17].

Minimi quadrati per il caso più generale di combinazioni lineari

Per il caso più generale di combinazioni lineari di funzioni di x si ha, [13]:

$$y^{mis} = \sum_k a_k \Phi_k(x) \quad (13)$$

Le funzioni ϕ base possono essere, p.e.:

$$1, x, x^2, \dots, x^{M-1} \quad (14)$$

Quindi y è un polinomio di grado $M-1$. Le funzioni base potrebbero essere costituite da seni e coseni, nel qual caso avremmo una serie armonica. Si può mostrare che la serie di Fourier è una curva dei minimi quadrati di una funzione del tipo:

$$J = \left\| y^{mis} - \sum_k a_k \Phi_k(x) \right\|^2 = \min \quad (15)$$

Problemi non lineari

Senza entrare in questo caso che richiede trattazione specifica, un modello nel caso di un problema non lineare è il seguente:

$$A(x) \cdot x - y^{mis} = 0 \quad (16)$$

Il sistema deve essere linearizzato, p.e. con l'algoritmo di Newton –Kantorovich.

Il vantaggio dei dati regolarizzati

Si è visto come la sperimentazione generi sempre problemi inversi che richiedono un modello matematico col quale regolarizzare i dati di misura.

Questo è il significato ingegneristico del principio ricordato nell'introduzione:

“Sia nell'ipotesi sia nell'esperimento la realtà è espressa matematicamente, in quanto ridotta ad un insieme di grandezze fisiche misurabili (non qualitative)...”

L'ingegneria non ricerca leggi rigorose. E' sufficiente una legge *pressappoco vera* basata su rilievi *pressappoco veri*, [18]. Queste leggi sperimentali sono sufficienti per una previsione ragionevole del comportamento del sistema,

“.....poiché ciò che interessa in ultima analisi non è la conoscenza dello stato di tensione, ma una attendibile previsione, sulla base di questa informazione, del comportamento dell'elemento strutturale....”. [19]

Nei termini più moderni del problema, tale previsione può essere fatta solo con un inquadramento statistico. Solo un modello statistico ricavato dalla sperimentazione consente la stima delle distribuzioni dei parametri critici e, come conseguenza, la stima dell'affidabilità strutturale del sistema, [20].

5. CONCLUSIONI

E' stata presentata l'applicazione della sperimentazione nelle varie fasi di sviluppo del processo di progettazione. Essa serve specialmente nella fase di chiarificazione del compito prima della progettazione concettuale per tradurre le esigenze del potenziale utilizzatore del prodotto in precisi requisiti tecnici.

Altri strumenti di valutazione dei punti più deboli o di selezione dei parametri più importanti sulla base dell'influenza che essi hanno sugli altri parametri del sistema, possono inoltre essere visti come modelli probabilistici e statistici di fenomeni empirici e quindi rientrano a buon diritto nei metodi sperimentali e, più in generale, nel metodo scientifico.

La sperimentazione sulle strutture introduce infine un altro aspetto interessante: i problemi sperimentali sono sempre problemi inversi, spesso mal-posti. Esistono soluzioni teoriche integrate da informazioni di natura euristica per risolvere queste difficoltà, ma il valore concettuale di essi risiede proprio nella loro particolare natura che prepara il progettista alla valutazione di problemi mal-strutturati nei quali le informazioni non sono mai sufficienti per determinare una soluzione. Essi costituiscono la regola nella progettazione.

Concludendo, la sperimentazione è una buona maestra per formare i progettisti. La pratica del laboratorio è concettualmente simile alla pratica della progettazione. La sperimentazione tuttavia ha una forte penalizzazione che oggi emerge drammaticamente:

“Modern science is fast-moving and no laboratory can exist for long with a program based on old facilities. Innovation and renewal are required to keep a laboratory on the frontiers of science and only if it remains on the frontiers will it have a long-term future. It is important that these resources be made available even in times of tight budgets. Starving the future to feed the present is a mistake - it leads to obsolescence and stagnation.....” [21].

BIBLIOGRAFIA

- [1] C. L. Dym, 1994. *Engineering Design*, Cambridge University Press
- [2] Report 1988, ABET. *Accreditation Board for Engineering and Technology*, N. Y.
- [3] Marcus Vitruvius Pollio, (70 a.C. – 15), *De Architectura*.
- [4] E. Severino, 1970. *Storia della Filosofia*- Rizzoli Ed.
- [5] G. Pahl & W. Beitz., 1999. *Konstruktion Lehre*, Springer Verlag, Berlin, Editor in inglese K. Wallace: 1997 *Engineering Design a Systematic Approach*, Springer Verlag.
- [6] ICSID, 2000. *International Council of Societies of Industrial Design* www.icsid.org.
- [7] R. Descartes, 1938. *Discorso del Metodo* (a cura di N. Abbagnano), A. Morano Editore, Napoli.
- [8] K. Keniston, 2003. “La crisi dell’algoritmo degli ingegneri”, Massachusetts Institute of Technology. Comunicazione privata.
- [9] A. Freddi, 2004. *Imparare a progettare: Principi e metodi del progetto concettuale*. (pp. 428). ISBN: 88-371-1512-1. Il testo è in corso di traduzione in lingua spagnola. Bologna: Pitagora Editrice (Italy).
- [10] D. Croccolo, L. Cristofolini, A. Freddi, M. Bandini, 2002. "Fatigue strength of shot-peened nitrided steel" *Fatigue Fract. Eng. M*, 25, pp. 695-707.
- [11] M. Comandini, 2008. “Progettazione con Metodologie Avanzate di Organi di Macchine sollecitati a Fatica, Corpi non Saintvenantiani” Tesi di Dottorato. DIEM, Fac. Ingegneria, Università di Bologna.
- [12] K.-H. Laermann, 2008. *Inverse Problems on Experimental Structural Mechanics*, Shaker Verlag Aachen, G.R.
- [13] J.F. Doyle, 2008. *Modern Experimental Stress Analysis* J. Wiley 2004; see also by the same Authors: *Hybrid Methods in: Experimental Mechanics Handbook*, Springer Verlag, 2008.

- [14] Liu X. Han, 2003. *Computational Inverse Techniques In Nondestructive Evaluation*, Crc Press, Boca Raton, London, New York. ISBN 0-8493-1523-9.
- [15] G. Olmi, A. Freddi, D. Croccolo, 2007. "In field measurement of forces and deformations at the rear part of a motorcycle and structural optimization", Proceedings of the 13th International Conference on Experimental Mechanics, Alexandroupolis, Greece, July 1-6, 2007. (vol. 1, pp. 435 - 436). ISBN: 978-1-4020-6238-4. DORDRECT: E.E. Gdoutos (NETHERLANDS).
- [16] A. Freddi, G. Olmi, D. Croccolo, 2002. "Ein neuer Aufnehmer zur direkten Messung der von einem Skifahrer auf die Skier Uebertragenen Kraefte und Momente" MTBMesstechnische Briefe, Ausgabe 1/2002, n. , pp. 3-10
- [17] C. Casavola C. Pappalettere, F. Tursi, "Non-uniform residual stress fields on sintered materials", 9th Youth Symposium on Experimental Solid Mechanics, Trieste, Italy, July 7-10, 2010, 132-137
- [18] E. Boncinelli, U. Bottazzini, 2000. *La serva padrona* Ed. Scienze e Idee
- [19] Prolusione del Prof. Lucio Lazzarino al III Convegno AIAS, 1975. Bologna
- [20] G. Olmi, A. Freddi 2010. "Fatica oligociclica su cappe e rotor di turboalternatori" Terza giornata di studio Ettore Funaioli , Bologna,-16 luglio 2009.
- [21] B. Richter, Premio Nobel in Fisica, 1976. *Autobiografia*, Sito Internet a suo nome.
- [22] W. Lidwell, K. Holden, J. Butler, 2010. "Universal Principles of Design" Rockport Publisher, Beverly, Massachusetts.

INFLUENCE OF DYNAMIC EFFECTS ON POINT EHL CONTACTS IN GEAR SYSTEMS

Marco Barbieri

*Dipartimento di Ingegneria Meccanica e Civile DIMeC,
Università di Modena e Reggio Emilia, Italy
E-mail: mark@unimore.it*

Francesco Pellicano

*Dipartimento di Ingegneria Meccanica e Civile DIMeC,
Università di Modena e Reggio Emilia, Italy
E-mail: francesco.pellicano@unimore.it*

Angelo Oreste Andrisano

*Dipartimento di Ingegneria Meccanica e Civile DIMeC,
Università di Modena e Reggio Emilia, Italy
E-mail: angelooreste.andrisano@unimore.it*

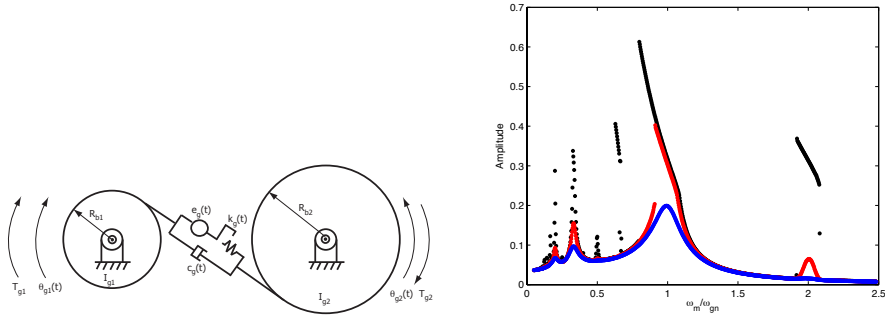
Abstract. *In this paper, the effect of vibrations on the elastohydrodynamical lubrication in spur gear pairs will be described. The relevance of inertial effects on film fluid lubrication is clarified by means of comparisons with static formulations. The multilevel technique is used to solve the transient EHL problem for elliptical contacts. Coupling with load balance equation and with different equations of motion is studied.*

A new model describing the dynamic behavior of two coupled transient EHL elliptical contacts is developed and applied to characterize the dynamics of a spur gear pair.

Keywords: *Gear dynamics, Transient EHL*

1. INTRODUCTION

Many models have been proposed in the past in order to describe the dynamic behavior of a spur gear pair, which is necessary to understand how vibration and noise is generated inside gearboxes. Many of these models are based on the work by Kahraman et al. [1], in which a one degree of freedom model is used to characterize the dynamic of the gear system. This model is depicted in Fig. 1(a). It includes a time varying stiffness, which is estimated by means of Finite Element Method static analyses, and a viscous damper, which value is determined only by means of experimental fitting. Some analyses performed on this model [2], showed that the results are strongly dependent on the dimensionless damping coefficient: Amplitude-Frequency diagram in Fig. 1(b) shows that for high viscous damping ($\zeta = 0.1$), the system behaves like a linear system, because there is no detachment of the contacting teeth. Due to this reason, it is important to clarify the influence of the lubricant in changing



(a) Lumped parameter model for spur gear dynamics (b) Amplitude-frequency diagram for different values of the non-dimensional damping coefficient (0.01 black, 0.05 red, 0.1 blue)

Figure 1: Model and results of conventional gear dynamic analysis

the damping effect.

The lubrication regime between mating gear teeth was proven to be elastohydrodynamic since the very earliest works on EHL, like those of Grubin [3]. Transient numerical analyses on gear lubrication were conducted by Larsson [4] and, more recently, by Wang et al. [5]. These works assume a given time dependent formula to describe the force transmitted by each mating tooth pair, either a square wave, for the first paper, or a trapezoidal one, for the second. Indeed, during dynamic meshing of the tooth pair, the contact force is strongly influenced by the inertial effects in the contacting bodies, as can be computed using the 1 dof proposed by Kaharaman, so that the results obtained using static approximations of the load are incorrect. The purpose of this paper is to propose a new model to describe the dynamic behavior of a spur gear pair including the effects of the lubrication.

2. TRANSIENT EHL FORMULATION

The problem of transient EHL will be solved for the point contact. This choice, although more computationally costly than studying line contacts, will allow to keep into account the effect of gear crowning and to extend the model to other kinds of gears. Transient Reynolds equation reads:

$$\frac{\partial}{\partial x} \left(\frac{\rho h^3}{12\eta} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(\frac{\rho h^3}{12\eta} \frac{\partial p}{\partial y} \right) - u_m \frac{\partial}{\partial x} (\rho h) - \frac{\partial}{\partial t} (\rho h) = 0 \quad (1)$$

where u_m is the mean velocity of the mating profiles and varies linearly during tooth contact. Film thickness is computed as follows:

$$h = h_0 + \frac{x^2}{2R_x} + \frac{y^2}{2R_y} + \frac{2}{\pi E'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{p(x', y', t) dx' dy'}{\sqrt{(x - x')^2 + (y - y')^2}} \quad (2)$$

where R_x is the equivalent radius of curvature in the profile direction and R_y along the face width. R_y is constant during tooth meshing, while R_x is varying with quadratic law. Figure 2

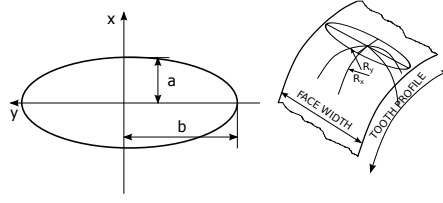


Figure 2: Contact ellipse and sketch of the equivalent tooth profile

explains the meaning of these parameters; note that equivalent radii are referred to a contact between a parabolic surface and a plane: the relationship among the true radii of curvature and the equivalent ones is the same as in standard Hertz's theory. Roelands pressure-viscosity relationship and Dowson and Higginson density-pressure formula are used to compute the dynamic viscosity η and density ρ . The equations are made dimensionless with respect to the hertzian point contact solution. Here, since many parameters change during meshing, a reference position for the gear system is to be chosen: the pitch point is the natural choice. Dimensionless P is defined dividing by the maximum hertzian pressure p_h at the pitch point, X and Y are divided by the minor half axis a^* of the hertzian contact ellipse, so the non-dimensionalization is as in the following:

$$\begin{aligned} X &= \frac{x}{a^*} \quad ; \quad Y = \frac{y}{a^*} \quad ; \quad H = \frac{R_x}{a^{*2}} h \quad ; \quad P = \frac{p}{p_h} \\ \bar{\eta} &= \frac{\eta}{\eta_0} \quad ; \quad \bar{\rho} = \frac{\rho}{\rho_0} \quad ; \quad U_m = \frac{t^*}{a^*} u_m \quad ; \quad T = \frac{t}{t^*} \end{aligned} \quad (3)$$

where t^* is to be defined so that the number of parameters is reduced to a minimum. The best choice is to obtain equal coefficients for the Couette and the Squeeze term inside the Reynolds equation. The resulting dimensionless equations are:

$$\begin{aligned} \frac{\partial}{\partial X} \left(\bar{\xi} \frac{\partial P}{\partial X} \right) + \frac{\partial}{\partial Y} \left(\bar{\xi} \frac{\partial P}{\partial Y} \right) - \frac{\partial}{\partial X} (\bar{\rho} H) - \frac{\partial}{\partial T} (\bar{\rho} H) &= 0 \\ \bar{\xi} = \frac{\bar{\rho} H^3}{\lambda \bar{\eta}} \quad \text{where} \quad \bar{\lambda} = \frac{12 u_m \eta_0 R_x^2}{a^{*3} p_h^*} \end{aligned} \quad (4)$$

$$H = H_0 + \frac{1}{R_x} \frac{X^2}{2} + D^* \frac{Y^2}{2} + \frac{2}{\pi^2} \chi_{p_h^*} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{p(X', Y', T) dX' dY'}{\sqrt{(X-X')^2 + (Y-Y')^2}}$$

where $\chi_{p_h^*}$ is a correction for the elliptical case and requires the numerical or approximated computation of elliptical integrals (see Ref. [2]). R_x is the ratio between the radius of curvature along the profile in a given position and the value at the pitch point; D^* is the ratio between R_x^* and R_y at the pitch point.

3. DYNAMIC MODEL

The model described in [1] has been modified here in order to include the effect of the EHL lubrication film. The contact between the two gears is shared by two lubricated contacts. Since the main purpose of this model is to describe damping, there is no need, at first, to consider the true deflection of the teeth under load, so the wheels in contact can be considered

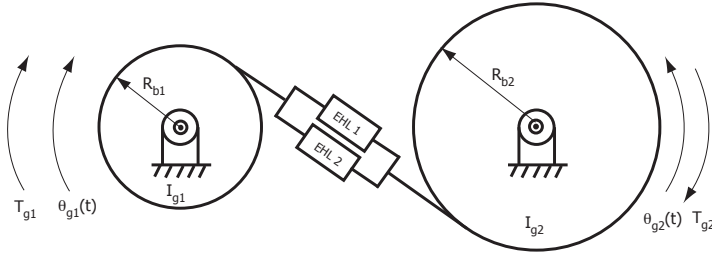


Figure 3: Full model for gear dynamics under EHL

as indefinitely extended elastic bodies. With such hypothesis, elastic deformations can be computed as in (2). It is to stress that the proposed model is an approximation of the physical system, but it is no more a lumped parameter model. The equation of motion is the following:

$$m_{ge} \frac{d^2 h_0}{dt^2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_1(x, y, t) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_2(x, y, t) dx dy + \frac{T_{g1}}{R_{b1}} \quad (5)$$

If two contacts share the load, it is possible to assume that the load they support is the same.

This way a single contact can be studied:

$$m_{ge} \frac{d^2 h_0}{dt^2} = \frac{1}{f(t)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y, t) dx dy + \frac{T_{g1}}{R_{b1}} \quad (6)$$

where $f(t)$ is a square wave function describing the time varying load on a single contact. In order to make both the equations non dimensional, it is useful to introduce the dimensionless inertia:

$$I = \frac{m_{ge}}{R_x^{*2} p_h^*} = \frac{2\pi m_e u_m^{*2}}{3R_x^* w^*} \frac{1}{\kappa^*} \quad (7)$$

Eq. (6) becomes:

$$I \frac{d^2 H_0}{dT^2} = \frac{1}{f(T)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(X', Y', T) dX' dY' - \frac{2\pi}{3\kappa^*} \quad (8)$$

4. RESULTS

Three different kinds of simulations are compared in the present section. The first analysis considers equation (6) with zero equivalent mass, so that the results are comparable to the literature [4] and [5]. Data of the gear pair under analysis are collected in Tab. ??: the analysis is performed on a spur gear pair, nevertheless the elliptical EHL model allows to keep into account the effect of crowning. The contact ratio of the gear pair is between 1 and 2, so that the maximum number of mating tooth pairs is two. The variation of the relative speed and of the equivalent curvature during meshing is taken into account. Figure 4(a) displays in blue the minimum H_m and central H_c dimensionless film thickness together with the central pressure P_c for the analysis without inertial effects. Close to the abrupt change of the load, there is a strong bump both in the central film thickness and in the central pressure; the minimum

Table 1: Data of the case study

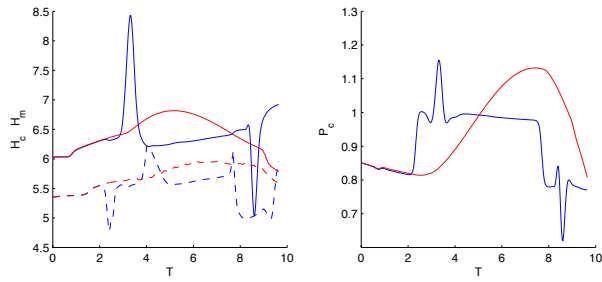
Data	Symbol	Pinion	Gear
Number of teeth	Z	28	43
Normal module [mm]	m	3.00	3.00
Pressure angle [Deg]	α	20.00	20.00
Base radius [mm]	R_b	39.467	60.610
Thickness at the pitch circle [mm]	s_p	6.1151	6.7128
Outer diameter [mm]	R_a	93.1	139.7
Crowning magnitude [mm]	mag_c	0.020	0.020
Face width [mm]	F_w	20.00	20.00
Inertia [$kg\ mm^2$]	I_g	8.076	47.762
Young's modulus [MPa]	E	206000	206000
Poisson's coefficient	ν	0.30	0.30
Center distance [mm]	a'_g	111.00	
Backlash on the line of action [mm]	j_n	0.312	
Transmission ratio	τ	0.6511	
Contact ratio	ε_α	1.2856	
Torque at pinion [Nmm]	T_{g1}	470000	

film thickness changes even faster, because the position of the minimum changes inside the computational domain. If the inertial effects are taken into account, here a simulation at a pinion speed of 12500 RPM is performed, the elastic oscillation is dominant and the squeeze effect is negligible to determine the dynamic behavior of the system. The main issues of the proposed approach are that the analysis cannot continue after the exit of a tooth pair contact and it is highly sensitive to the initial conditions: here stationary EHL film fluid at the entering is assumed, which is pretty unphysical, as the film fluid needs to build up before having true stationary conditions.

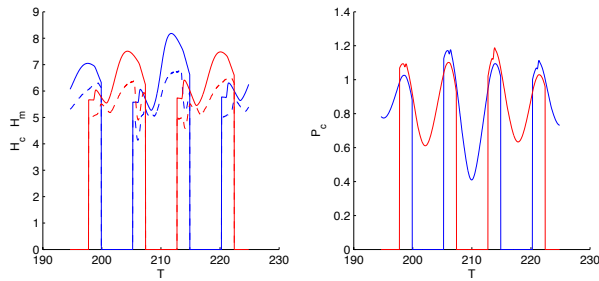
In order to overcome this limitation, an analysis considering the coupling between two EHL contacts is performed. Initial conditions are set up assuming stationary EHL film at the pitch point for one contact. Since at the pitch point only one pair is in contact, the analysis on the second contact starts with the same approach h_0 as the first contact, which is still active up to the exit. The solution oscillates due to the initial condition, so some transient mesh cycles must be thrown in order to have a stationary solution. Figure 4(b) shows the behavior after 200 mesh periods: the entering of each contact produces some oscillation in the film thickness as it was for the static case, but they are smoothed by the inertia in the contacting bodies.

5. CONCLUSIONS

A new method to describe the Elastohydrodynamic lubrication in a spur gear pair has been proposed. The method takes into account the dynamic load inside a single tooth pair and the coupling between two different pairs. Results on a real case study are shown: if the dynamic load is neglected, there is a fast variation both in the pressure distribution and in the film thickness when the number of mating tooth pair changes. If the inertia of the contacting bodies is considered, as is the present work, the transition is by far smoother.



(a) Single contact: blue - static analysis; red - dynamic analysis



(b) Full model: blue - contact 1; red - contact 2

Figure 4: Results of the EHL-based models

REFERENCES

- [1] Kahraman, A., Singh, R., 1990, "Non-linear Dynamics of a Spur Gear Pair", *Journal of Sound and Vibrations*, **142** (1), pp. 49-75.
- [2] Barbieri, M., 2010, "Gear Dynamics Under EHL Conditions", PhD Thesis, Università di Modena e Reggio Emilia, Modena, Italy.
- [3] Grubin A.N., 1949, *Fundamentals of the Hydrodynamic Theory of Lubrication of Heavily Loaded Cylindrical Surfaces*, Moscow: TsNIITMASH, Book 30 (Engl. transl.).
- [4] Larsson, R., 1997, "Transient non-Newtonian Elastohydrodynamic Lubrication Analysis of an Involute Spur Gear", *Wear*, **207**, 67-73.
- [5] Wang, T., Li H., Tong, J., Yang, P., 2004, "Transient Thermoelastohydrodynamic Lubrication Analysis of an Involute Spur Gear", *Tribology International*, **37**.

FATICA TERMOMECCANICA DI FILI A MEMORIA DI FORMA: APPARATO DI PROVA E PRIMI RISULTATI

Giovanni Scire' Mammano

Department of Scienze e Metodi dell'Ingegneria,

University of Modena and Reggio Emilia, Italy

E-mail: giovanni.sciremammano@unimore.it

Eugenio Dragoni

Department of Scienze e Metodi dell'Ingegneria,

University of Modena and Reggio Emilia, Italy

E-mail: eugenio.dragoni@unimore.it

Sommario. Il lavoro riguarda la fatica termomeccanica, nota anche come fatica funzionale, di fili commerciali a memoria di forma. Da un'analisi della letteratura disponibile emerge come la caratterizzazione a fatica di questi fili si limiti quasi esclusivamente alla condizione di isotensione, peraltro poco riconducibile alle reali condizioni di utilizzo dei fili negli attuatori. La presente ricerca cerca di colmare questa lacuna individuando 4 tipi di caratterizzazione più aderenti alla realtà applicativa: prove in isotensione, isodeformazione, isotensione con deformazione limitata, ciclo di tensione con deformazione limitata. Nel presente lavoro si espongono nel dettaglio le modalità di prova e la macchina messa a punto per la loro implementazione sperimentale. Infine si presentano i risultati delle prove in isotensione, prima fase della campagna sperimentale.

Parole chiave: fili SMA, fatica termomeccanica, fatica funzionale, isotensione, isodeformazione

1. INTRODUZIONE

I materiali a memoria di forma (SMA = Shape Memory Alloys) sono sempre più utilizzati per la realizzazione di semplici attuatori allo stato solido, caratterizzati dall'elevata densità di potenza. La progettazione di questi attuatori richiede la conoscenza del comportamento a fatica della lega SMA utilizzata [1]. Negli attuatori l'attivazione del materiale è sempre di tipo termico, per cui si parla in genere di fatica termomeccanica, chiamata anche fatica funzionale. Con questo termine si indica il funzionamento del materiale sotto l'azione combinata di un ciclo termico e di uno meccanico. Il primo ciclo indispensabile per attivare la lega provoca l'insorgere di un ciclo di lavoro meccanico dovuto al carico applicato all'elemento SMA.

In letteratura, il fenomeno della fatica funzionale non è trattato estesamente come quello della fatica meccanica classica, per la quale sono a disposizione un vasto numero di pubblicazioni. Il motivo principale è che ad oggi l'utilizzo dei materiali a memoria di forma, ed in particolare del Nitinol, è molto sbilanciato su applicazioni che sfruttano le proprietà superelastiche e non l'effetto memoria di forma. Inoltre i pochi lavori disponibili, pubblicati sia da enti di ricerca che da aziende produttrici, sono limitati principalmente alla

condizione di tensione costante applicata al materiale (Isotensione) e presentano risultati di prove con numero di cicli limitato. In [2-3] viene effettuata una caratterizzazione a fatica di fili commerciali SAES Getters® SmartFlex®. La modalità di prova proposta prevede di caricare il filo in condizioni di isotensione e di variare la deflessione recuperata interrompendo l'alimentazione elettrica al filo al raggiungimento della corsa voluta (trasformazione parziale). Il risultato sintetico delle prove [2] è un grafico deformazione-tensione sul quale sono indicate zone di funzionamento caratterizzate da una vita minima garantita. In [4] viene analizzato il comportamento a fatica a tensione costante di fili a memoria di forma analizzando anche l'interazione di un'eventuale corrosione superficiale dei fili stessi. In [5] Eggeler sottopone delle molle a memoria di forma a prove di fatica a carico costante, evidenziando come all'accumularsi del numero di cicli vi sia un allungamento permanente ma al contempo un aumento della corsa prodotta. Lagoudas [6] caratterizza il comportamento a fatica termomeccanica di fili SMA NiTiCu. In particolare studia l'influenza del trattamento termico di ricottura (annealing) sulla resistenza a fatica del materiale e individua un trattamento ottimale. Nel lavoro si riportano anche i risultati dell'influenza, sulla resistenza a fatica, della completa o parziale trasformazione di fase indotta dal riscaldamento. Secondo gli autori, i principali fattori che influenzano il comportamento a fatica sono la tensione meccanica applicata, il grado di trasformazione e il processo subito dal materiale. In [7] Bignon e Morin confrontano il comportamento a fatica di due leghe commerciali: TiNi e CuZnAl. Nel lavoro si mette in evidenza come l'effetto di memoria di forma sia influenzato sia dal carico applicato che dall'accumularsi del numero di cicli. In particolare questa seconda dipendenza ha un effetto benefico (training) ed è maggiormente significativo per la lega CuZnAl. In [8] i due autori riprendono l'analisi della fatica termomeccanica sulle due leghe. Dal lavoro emerge come esista una relazione diretta tra SATWME (Stress Assisted Two Way Memory Effect) e carico applicato e che il massimo effetto di recupero di forma si ottiene, nel caso di lega TiNi per un carico applicato di circa 200MPa. In [9] uno studio analogo è proposto per fili in TiNiCu per i quali si presenta anche una curva di Wohler per la previsione della vita.

Altre tipologie di prova, differenti da quella in isotensione sono proposte in [10] e [11]. Mertmann et al. [10] studiano la stabilità dell'effetto memoria di forma nel caso di recupero di forma vincolato, ossia nel caso di isodeformazione. Dalle prove emerge la scarsa attitudine del materiale a lavorare in questa modalità che si traduce all'accumularsi dei cicli effettuati in perdita della memoria ("amnesia") e allungamento del provino. Demers et al. [11] studiano la resistenza a fatica termomeccanica di una lega Ni-Ti, al variare di differenti trattamenti termomeccanici di allenamento. Per la valutazione dell'efficacia dei trattamenti suggeriscono un'interessante metodologia che si articola su tre differenti prove: recupero libero ("Stress-free shape recovery"), recupero vincolato o isodeformazione ("Constrained stress recovery") e recupero a carico costante o isotensione ("Assisted two-way shape memory effect ATWSME").

La difficoltà che il progettista incontra nel dimensionamento di un dispositivo basato su elementi a memoria di forma è dovuta principalmente a due motivi. In primo luogo la mancanza di dati direttamente utilizzabili sulla resistenza a fatica dello specifico materiale commerciale con cui si realizzerà il dispositivo, come ad esempio curve di Wohler. In secondo luogo la scarsa affinità delle prove isotensione, pressoché le uniche disponibili come visto, con il reale funzionamento degli attuatori SMA.

Gli attuatori industriali sono infatti quasi sempre contrastati da elementi elastici convenzionali (molle) o da altri elementi SMA contrapposti. In entrambi i casi la

condizione di carico a cui l'elemento SMA è sottoposto è molto differente rispetto alla condizione di tensione costante realizzata nelle prove di laboratorio e questo porta a modelli di attuatore non ottimali e alla pressoché impossibilità di previsione della vita. Questo lavoro si propone di colmare il divario tra sperimentazione e realtà. In quest'ottica sono state ideate 4 modalità di prova differenti che si avvicinano alla maggior parte delle reali situazioni di funzionamento dei fili SMA: carico costante (isotensione), deformazione costante (isodeformazione), carico costante con deformazione limitata, ciclo di tensione con deformazione limitata. Il filo utilizzato nelle prove è un filo commerciale SAES Getters[®], tipo Smartflex[®] 150, di diametro 0.15mm, consigliato per la realizzazione di attuatori. Nell'articolo si presenta l'apparato sperimentale progettato specificamente per attuare le quattro condizioni di prova. Inoltre si presentano i risultati completi delle prove in isotensione.

2. MATERIALI E METODO

Tipologie di prova

Come ben noto un attuatore a memoria di forma è costituito da un elemento attivo, ad esempio un filo, e da un sistema di contrasto capace di deformarlo quando questo si trova nello stato disattivato. Il carico a cui l'elemento è sottoposto dipende quindi dal sistema di contrasto e dal carico esterno ad esso applicato. Se si esclude il caso di contrasto mediante grave e carico esterno costante o assente, la sollecitazione a cui l'elemento attivo è sottoposto non è costante durante il ciclo di attivazione-disattivazione, ma variabile tra un massimo e minimo, da valutare caso per caso.

Qualora l'attuatore sia dotato di finecorsa meccanici, il carico risulta essere ancora variabile, ma la deformazione a cui è sottoposto il filo risulta limitata inferiormente, superiormente, o in entrambi i versi. Ultimo caso è quello in cui l'elemento attivo compia la trasformazione essendo vincolato: in questo caso la tensione è ancora variabile ma la deformazione è fissa.

Diversamente da quanto riportato in [11], il caso di recupero di forma vincolato non si verifica solo per attuatori funzionanti come generatori di forza statici. Il recupero vincolato può infatti verificarsi per qualsiasi attuatore qualora sia utilizzato per movimentare o organi di presa o carichi fortemente inerziali con dinamiche elevate. Nel primo caso l'evento si verifica qualora l'organo di presa afferri un oggetto di dimensioni al limite delle sue capacità [10]. Il secondo caso si verifica invece quando la costante di tempo di riscaldamento dell'elemento SMA sia inferiore alla costante di tempo del carico meccanico applicato all'attuatore, dipendente dalla sua inerzia. Dal punto di vista pratico questo avviene tutte le volte che si riscalda molto rapidamente il materiale SMA [12] e/o l'inerzia del carico è elevata. Di fatto il filo completa la sua trasformazione sotto l'azione di un vincolo virtuale che non ne permette la deformazione. Infine non va dimenticata la condizione di bloccaggio dell'attuatore per cause anche esterne dell'attuatore stesso. Anche in tali situazioni si verifica un recupero di forma vincolato ed è auspicabile che l'attuatore sia intrinsecamente robusto di fronte ad eventi di questo tipo.

Alla luce delle condizioni di sollecitazioni dei fili SMA nella realizzazione di attuatori si propongono quattro modalità di prova utili ad ottenere informazioni più vicine alle reali condizioni di utilizzo.

La prima modalità di prova è la classica prova in "Isotensione", in cui al filo SMA è applicato un carico gravitazionale costante. La curva di lavoro del filo nel piano tensione-

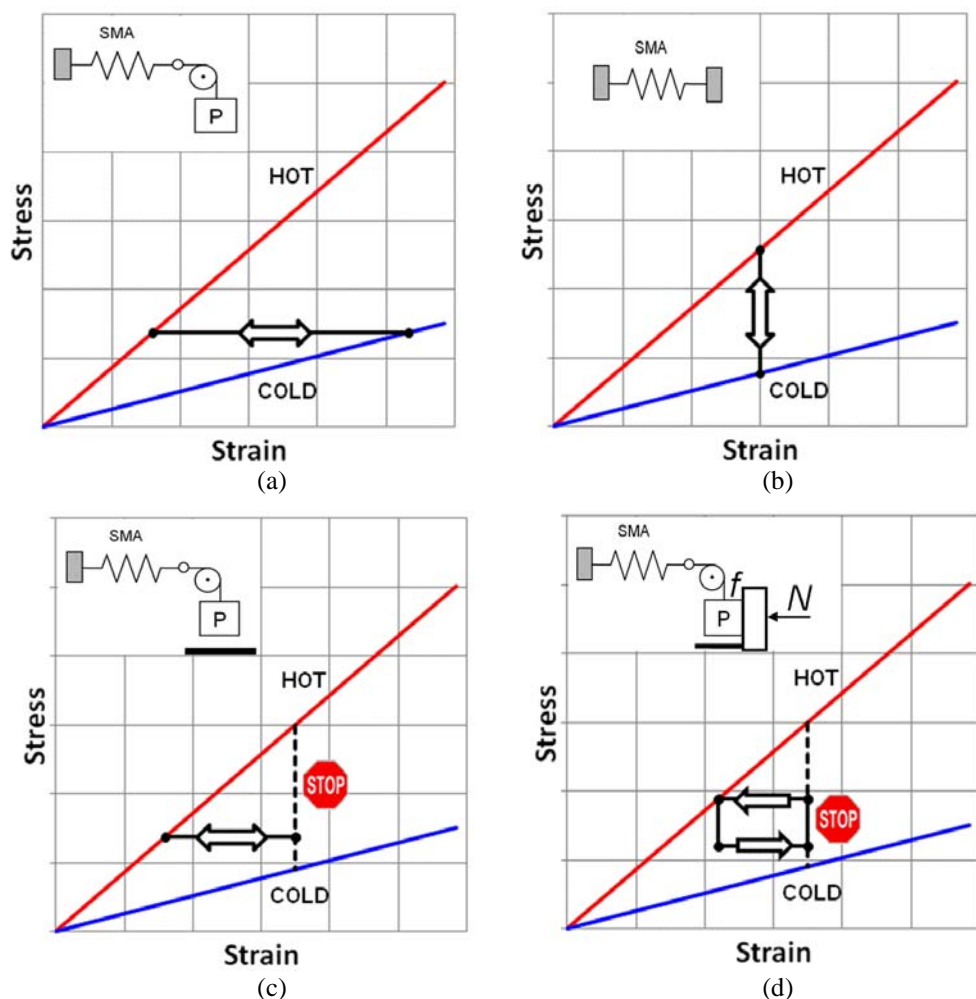


Figura 1. Ciclo tensione-deformazione applicato al filo nelle differenti modalità di prova proposte: a) Isotensione , b) Iso deformazione, c) Isotensione con deformazione limitata, d) ciclo di tensione con deformazione limitata

deformazione è una retta orizzontale (Fig. 1a). L'implementazione di questa modalità di prova è molto semplice in quanto è sufficiente contrastare il filo con un grave di massa opportuna. La prova risulta caratterizzata da un solo parametro: la tensione nel filo.

La seconda modalità è la prova in "Iso deformazione". In questa modalità il filo effettua la trasformazione cristallina senza poter variare la sua forma. La curva di lavoro nel piano tensione-deformazione è una retta verticale (Fig. 1b). Per realizzare questa prova è sufficiente vincolare il filo ad ambo gli estremi imponendogli un allungamento tale da realizzare il livello di deformazione voluta. Anche in questa prova si agisce su un solo parametro: la deformazione del filo. La terza modalità di prova è la prova in "Isotensione con deformazione limitata": questa prova è ancora una prova in isotensione, ma

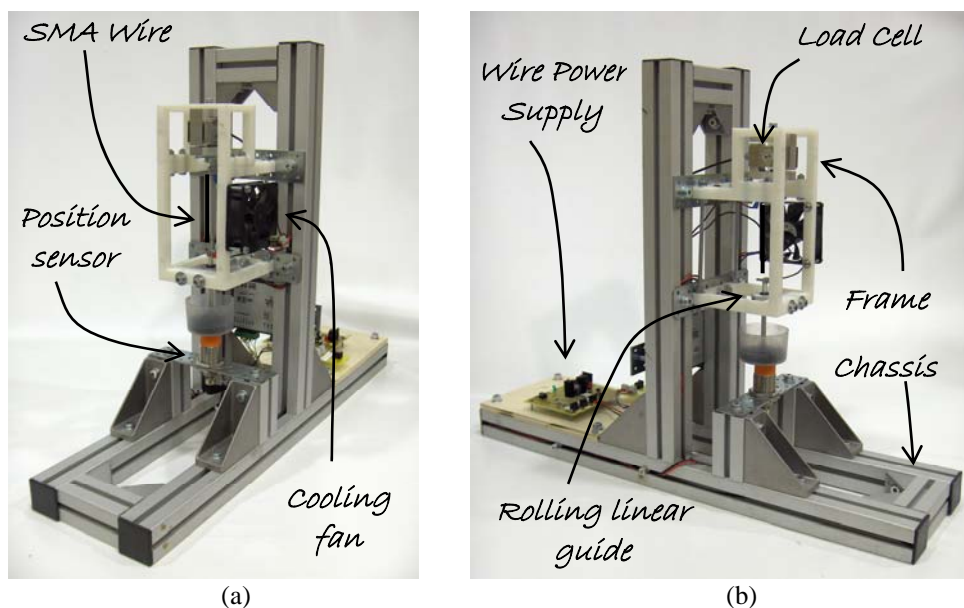


Figura 2. Macchina realizzata per l'esecuzione delle prove di fatica su fili SMA secondo le 4 modalità di prova proposte: a) lato destro, b) lato sinistro

caratterizzata dal fatto che la deformazione massima che il filo può raggiungere è limitata ad un valore prefissato. In questo modo è possibile far lavorare il filo a tensioni molto maggiori di quelle implementabili con la prova in isotensione classica, senza al contempo raggiungere deformazioni che porterebbero il filo a danneggiarsi o a rompersi in pochi cicli. La curva di lavoro nel piano tensione-deformazione è ancora una retta orizzontale, ma limitata a destra al valore di deformazione massima (Fig. 1c). Questa tipologia di prova fa sperimentare al filo un ciclo di lavoro più simile a quello che avrebbe nel caso di utilizzo in un attuttore genericamente contrastato e dotato di finecorsa meccanici. La prova è caratterizzata da due parametri: la tensione applicata e la deformazione massima. Per implementare questa prova è sufficiente aggiungere all'apparato della prova in isotensione un sistema che consenta di limitare lo spostamento massimo del grave. L'ultima modalità di prova è invece la prova a "Ciclo di tensione con deformazione limitata". In questa prova il filo sperimenta due differenti valori di tensione: durante la fase di recupero della forma il filo deve vincere una tensione maggiore della tensione che lo deforma nella successiva fase di deformazione. Inoltre la deformazione massima che il filo può raggiungere è ancora limitata ad un valore prefissato. La curva di lavoro risultante è un rettangolo (Fig. 1d). Questa tipologia di prova consente di implementare nei fili dei cicli di lavoro molto simili a quelli che sperimentano negli attuatori, dove i carichi esterni (non conservativi) che l'attuttore deve vincere contrastano sia la fase di recupero che la fase di deformazione. Una possibile realizzazione di questa modalità di prova prevede di affiancare al grave, nell'azione di contrasto del filo, un elemento dissipativo ad attrito. Questo elemento genererà una forza in direzione opposta al moto del grave stesso che quindi si sommerà al

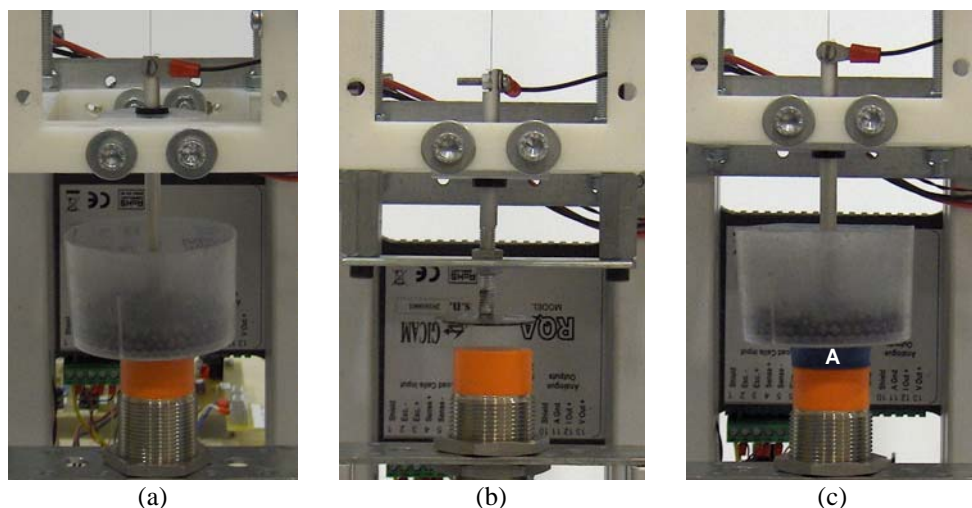


Figura 3. Accessori per l'esecuzione delle prove in: a) Isotensione, b) Isodeformazione, c) Isotensione con deformazione limitata

peso nella fase di recupero di forma e si sottrarrà nella fase di deformazione. Variando in modo oculato l'intensità della forza di attrito e la massa del grave è possibile generare qualsiasi combinazioni dei due livelli di tensione. Analogamente alla prova in isotensione con deformazione limitata deve essere presente un apparato che consenta di limitare lo spostamento massimo del grave. Questa prova è caratterizzata da tre parametri: i due livelli di tensione del ciclo e la deformazione massima.

Apparato sperimentale

Al fine di implementare le quattro modalità di prove precedentemente esposte è stata progettata e realizzata una macchina per la caratterizzazione a fatica di fili a memoria di forma. La macchina, visibile in Fig. 2a e 2b, è costituita da un telaio a traliccio in alluminio. Alle colonne verticali del telaio è fissato un telaio in materiale plastico che accoglie sulla parte superiore una cella di carico, alla quale è possibile fissare, mediante un serraggio a vite, un estremo del filo SMA in prova. La cella di carico ha lo scopo di misurare la forza esercitata dal filo. Il banco è predisposto per l'allestimento di due differenti celle diversificate per portata: una cella per prove a basso carico (isotensione) di portata 3kgf e una per prove ad alto carico (isodeformazione) di portata 25kgf.

L'altro estremo del filo è collegato sempre mediante un serraggio a vite ad un perno (differenziato a seconda del tipo di prova), guidato verticalmente da una guida lineare a ricircolo di sfere. Il perno presenta nell'estremità inferiore un disco ferromagnetico che si affaccia sulla superficie di misura di un sensore di posizione induttivo, capace di misurare gli spostamenti del perno stesso fino ad una distanza massima di 15mm.

Il raffreddamento del filo a convenzione forzata è garantito da una ventola di raffreddamento, il cui numero di giri è regolabile mediante l'utilizzo di un alimentatore da laboratorio. Nella configurazione attuale la macchina è predisposta per l'esecuzione delle prove in isotensione, isodeformazione e isotensione con deformazione limitata.

Nel caso di prove in isotensione si utilizza il sistema visibile in Fig. 3a, costituito da un perno liscio in alluminio recante all'estremità il disco in materiale ferromagnetico indispensabile per il rilevamento della posizione dell'apparato mobile da parte del sensore di posizione. Il disco stesso funge inoltre da supporto per il cestello portapesi. Per garantire un'ottima risoluzione del carico desiderato, il cestello viene riempito con pesetti di piombo di massa 0.3 grammi. La massa del perno e del cestello vuoto, corrispondente al minimo carico applicabile, è di soli 15 grammi.

Per l'esecuzione delle prove in isodeformazione invece è necessario smontare il perno liscio con cestello e montare il sistema di pretensionamento del filo illustrato in Fig. 3b. Il sistema è costituito da un perno parzialmente filettato inserito con gioco nel foro centrale di una piastrina orizzontale in acciaio, vincolata alla frame mediante due viti. Sul perno sono inseriti due dadi esagonali ribassati, uno sopra e uno sotto la piastrina. Regolando la posizione del dado inferiore è possibile regolare la deflessione imposta al filo, che viene quindi congelata mediante serraggio del controdado superiore. Sfruttando la lettura della posizione fornita dal sensore induttivo è possibile regolare con ottima precisione l'entità della deflessione del filo e conseguentemente, nota la lunghezza del filo stesso, la deformazione imposta. Il peso del perno e del disco ferromagnetico per le prove in isotensione è di 18 grammi.

Nel caso di prove in isotensione a deformazione limitata, la macchina viene accessoriata con il sistema della prova in isotensione classica. L'unica differenza risiede nell'aggiungere un elemento distanziatore fra la testa di misura del sensore e il disco ferromagnetico posto all'estremità del perno. Il distanziatore "A" visibile in Fig. 3c è costituito da un disco in materiale elastomerico ad alta durezza, ed è appoggiato sulla superficie di misura del sensore induttivo. Essendo in materiale polimerico, il disco non viene rilevato e non altera la misura del sensore. Il suo scopo è duplice: in primo luogo ha la funzione di bloccare la discesa del cestello, limitando quindi la deformazione massima del filo; in secondo luogo la sua interposizione fra sensore e disco consente di assicurare sempre la distanza minima sensore-disco ferromagnetico indispensabile per ottenere una risposta lineare del sensore stesso. Noto lo spessore del disco elastomerico è possibile regolare l'entità della deformazione massima agendo sulla posizione del sensore e sfruttando la lettura di posizione del sensore.

Una scheda elettronica appositamente realizzata fornisce l'alimentazione del filo SMA. La scheda sfrutta un regolatore lineare di tensione (Linear LT3080) retroazionato sulla corrente che circola sul filo. Il riferimento di corrente è fornito alla scheda mediante un segnale in tensione proporzionale all'intensità di corrente desiderata. Tutti i sensori e trasduttori con cui la macchina è equipaggiata (cella di carico, sensore di posizione, sensore di corrente sul filo SMA, sensore temperatura aria di raffreddamento) sono acquisiti mediante una scheda DAQ della National Instruments® (USB 6251) e visualizzati e memorizzati su personal computer grazie ad un applicativo appositamente realizzato in ambiente LabView®. L'applicativo, congiuntamente alla scheda DAQ, si occupa anche alla generazione del segnale di tensione che comanda la corrente erogata dalla scheda elettronica. In Fig. 4 è visibile l'interfaccia di comando LabView® della macchina realizzata. Il programma ha bisogno dei dati essenziali del filo, come diametro e lunghezza allo stato indeformato oltre al tipo e alle caratteristiche del ciclo di corrente che si vuole imporre al filo. Tra i tipi di cicli è possibile scegliere fra forma d'onda sinusoidale, triangolare, dente di sega, quadra. Relativamente ai parametri è possibile fissare il valor medio dell'onda, l'ampiezza, la frequenza e nel caso di onda quadra il duty cycle.

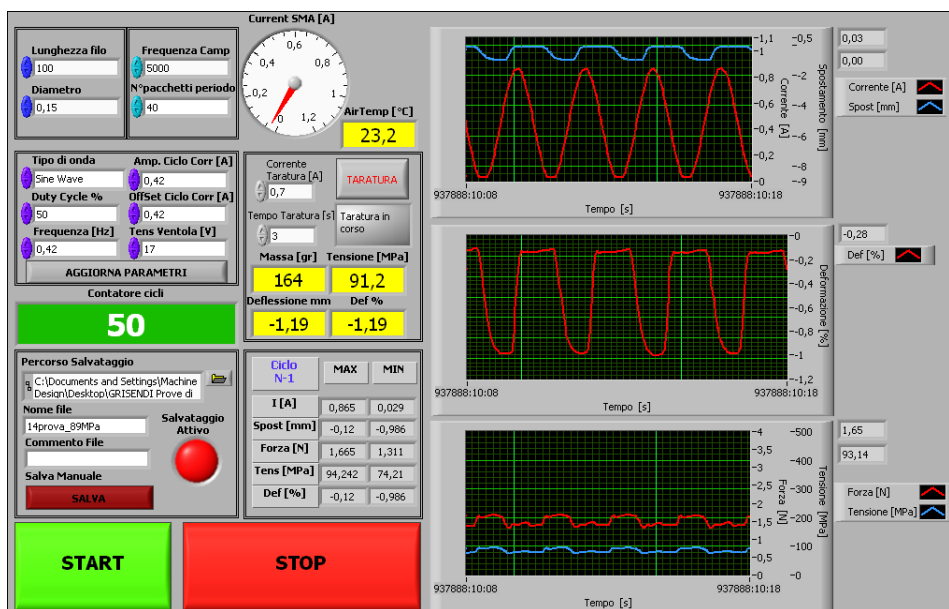


Figura 4. Interfaccia del programma di controllo e gestione della macchina di prova

Riguardo alla memorizzazione dei dati sperimentali il programma esegue due politiche di salvataggio. In primo luogo memorizza su un file, ad ogni ciclo, i valori massimi e minimi delle grandezze acquisite. Inoltre ad intervalli progressivi (ciclo 10, 20, 30, ..., 100, 200, 300, ..., 1000, 2000, ...) salva per un ciclo completo del filo tutte le grandezze acquisite ad una frequenza impostabile dall'operatore, solitamente fissata a 5kHz. In tempo reale poi il programma permette di vedere i tracciati di forza, tensione meccanica, spostamento, deformazione e corrente elettrica, oltre a visualizzare i valori di massimo e minimo delle grandezze al ciclo precedente.

L'applicativo si occupa anche dell'azzeramento dei sensori a inizio prova e dell'ausilio all'operatore nelle fasi di allestimento della prova (caricamento e posizionamento).

Protocollo della prova isotensione

Il protocollo messo a punto per la corretta esecuzione delle prove in isotensione prevede tre fasi.

Nella prima fase si stabiliscono, mediante prove preliminari, i parametri del ciclo di alimentazione elettrica del filo SMA e l'entità del raffreddamento forzato del filo. L'obiettivo è quello massimizzare la frequenza di ciclo e conseguentemente ridurre il tempo di prova, garantendo al contempo la correttezza della prova. In particolare è necessario garantire che le trasformazioni martensite – austenite e viceversa siano complete, che siano evitati surriscaldamenti del filo e che la tensione meccanica a cui il filo è soggetto sia costante durante l'intero ciclo. Sebbene infatti il filo sia caricato con un peso costante, la tensione subisce inevitabili variazioni a causa sia dell'attrito della guida lineare che delle forze di inerzia dovute alle accelerazioni e decelerazioni del cestello. Il protocollo prevede che l'oscillazione di tensione reale non superi il 5% della tensione nominale

Tabella 1. Caratteristiche del filo Saes Smartflex® (Dati forniti da SAES Getters)

Prova isteresi @ 200MPa, 1°C/min	
Stroke	4.8%
A_s	86°C
A_f	94°C
M_s	65°C
M_f	57°C
Fatica @ 3.5% 170MPa 0.6A	
N° Cicli	>100.000
Deriva	0.17%

applicata. Tutte le prove del presente studio sono state eseguite imponendo al filo un ciclo di corrente elettrica di tipo sinusoidale con valore minimo nullo e conseguentemente valore medio della sinusoide pari alla metà della corrente massima.

La seconda fase prevede l'allestimento e la preparazione della macchina per la specifica prova. Innanzi tutto, uno degli estremi del filo SMA in prova viene fissato ad uno dei due afferraggi. Successivamente si riscalda il filo per effetto Joule fino alla completa trasformazione del materiale allo stato austenitico, in modo da essere certi che il filo sia alla lunghezza memorizzata. Quindi, dopo aver serrato anche l'altro estremo, si riscalda nuovamente il filo in stato austenitico e in questa configurazione si esegue l'azzeramento del sensore di posizione e del sensore di forza. Al termine dell'azzeramento, il software LabView aggiunge in automatico al segnale di forza della cella di carico il peso noto del perno-cestello. Il programma inizia quindi la visualizzazione in tempo reale della grandezze acquisite, assecondando l'operatore nell'azione di caricamento del cestello con i pesetti di piombo fino al raggiungimento del livello di tensione meccanica σ stabilito. Terminato il caricamento del cestello l'operatore può avviare la terza ed ultima fase del protocollo ossia la prova in isotensione. A questo punto la macchina è completamente indipendente e fa ciclare il filo SMA fino al raggiungimento della rottura del filo, evento rilevato in automatico dalla macchina mediante un controllo sulla corrente realmente circolante e comunicato all'operatore mediante un messaggio di posta elettronica. In alternativa, la prova può terminare a seguito di un intervento manuale dell'operatore, che può in qualsiasi momento interrompere l'esperimento.

Materiale SMA utilizzato

L'obiettivo primario della presente attività di ricerca è quello di investigare la resistenza a fatica delle leghe SMA idonee allo sviluppo di attuatori. In particolare l'attenzione è rivolta alle leghe già disponibili sul mercato e quindi direttamente utilizzabili dalle imprese per lo sviluppo di applicazioni concrete. La quasi totalità della letteratura sulla fatica delle leghe SMA è invece basata su prove relative a leghe sperimentali, o per composizione o per trattamenti termomeccanici di allenamento.

Con l'intento di valutare differenti fili di diversi produttori il protagonista di questa campagna di prove è un filo della SAES Getters®, indicato dalla stessa azienda come idoneo allo sviluppo di attuatori. Il filo, di diametro 0.150mm, ha denominazione commerciale Smartflex® 150, NiTi/54/W/0.150/T150-01.

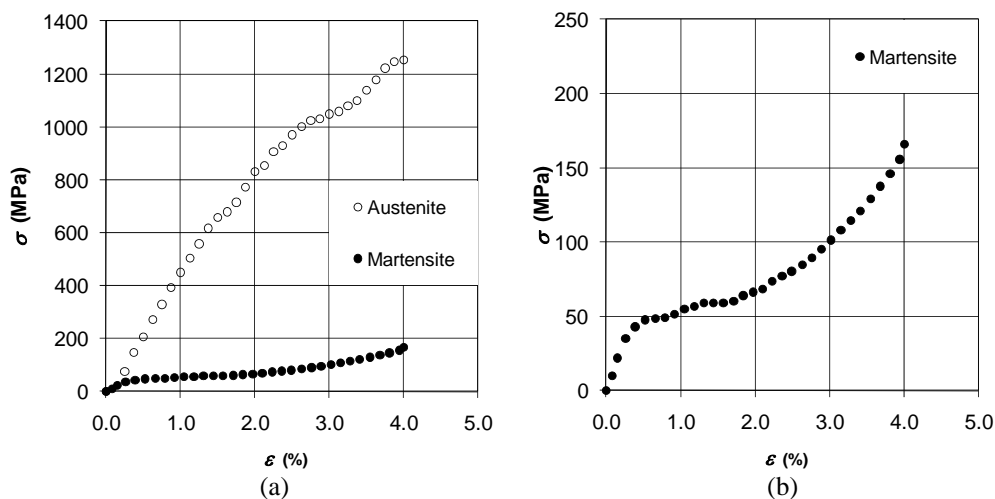


Figura 5. Diagramma sperimentale di trazione del filo Smartflex® di Saes Getters da 0.150mm utilizzato nelle prove: a) fase austenitica (riscaldamento a 0.5A) e martensitica b) Ingrandimento del comportamento in stato martensitico

Le caratteristiche fisico-meccaniche fornite dal produttore sono riassunte in Tab 1. In Fig. 5a si riporta il diagramma tensione – deformazione del materiale. La curva superiore (pallini vuoti) è riferita allo stato austenitico ($T > A_f$), ottenuto mediante riscaldamento per effetto Joule in convezione naturale con una corrente di alimentazione di 500mA. La curva inferiore (pallini pieni) è invece relativa allo stato martensitico ($T < M_f$) ed è visibile anche in scala ingrandita in Fig. 5b. Le prove sono state condotte con un dinamometro elettromeccanico Galdabini SUN 500 ad una velocità di deformazione di $1.67 \cdot 10^{-4} \text{ s}^{-1}$.

Pianificazione prove sperimentali isotensione

Obiettivo primario delle prove in isotensione è quello di ricavare la curva di Wohler σ -N del materiale. La procedura di prova adottata è aderente allo standard giapponese “Standard Method of Statistical Fatigue Testing, JSME 002-1981” [13]. Questo metodo permette di definire i livelli di tensione e fornisce i metodi per l’elaborazione statistica dei risultati delle prove di fatica. Il metodo identifica la curva di Wohler del materiale mediante 14 prove, 8 utilizzate per la determinazione della parte inclinata e 6 per la parte orizzontale della curva (limite di fatica). I dati ottenuti dalle prove relative alla parte inclinata della curva sono utilizzati per ottenere il corretto passo delle successive prove per la valutazione della parte orizzontale della curva, basate su un metodo “staircase” ridotto. Il metodo JSME suggerisce di eseguire alcune prove preliminari di fatica utili per definire i livelli di tensione da applicare. In particolare richiede di conoscere la tensione meccanica che porta alla durata di 5×10^4 e 1×10^7 . A seguito di alcune prove preliminari sul filo SMA in esame, queste durate sono risultate troppo elevate per il materiale. Per questo motivo si è deciso di utilizzare il metodo ponendo come durata inferiore e come durata limite di fatica i valori 5×10^3 e 5×10^5 , per i quali i livelli di tensione sono risultati 175MPa e 100MPa rispettivamente.

Tabella 2. Livelli di tensione applicata al filo stabiliti secondo il metodo JSME

Livello	Tensione (MPa)
S ₀ +3	225
S ₀ +2	200
S ₀ +1	175
S ₀	150
S ₀ -1	125
S ₀ -2	100
S ₀ -3	75

Noti questi valori il metodo fornisce una serie di livelli di prova riportati in Tab. 2. Le prove cominciano dal livello S₀ e proseguono scendendo di livello fino a quando si arriva ad una non rottura. Successivamente si risale di livello fino a completare il piano di prove in modo da ottenere 2 ripetizioni per ciascuno dei 4 livelli superiori a quello in cui si è registrata la rottura. Come detto si è deciso di considerare come non rottura una prova che superasse i 5×10^5 cicli.

Terminate le prove relative al tratto inclinato, il metodo utilizza la deviazione standard del tratto inclinato come passo della successiva procedura di staircase per l'individuazione del limite di fatica del materiale. Il metodo è strutturato in modo tale che il livello di confidenza dei due tratti della curva di Wohler, tratto inclinato e tratto orizzontale, sia il 50%. Dai risultati ottenuti nelle prove di fatica relative al tratto inclinato è risultato un passo di staircase di 11MPa. In Tab. 3 sono riassunte in ordine cronologico, tutte le 14 prove effettuate per la determinazione della curva di Wohler secondo il metodo JSME, completo dell'impostazioni utilizzate per l'alimentazione del filo SMA.

Tabella 3. Successione delle 14 prove in isotensione con indicazione dei parametri di alimentazione del filo

Prova	Livello	Tensione (MPa)	Frequenza (Hz)	I _{max} (A)
1	S ₀	150	0.35	1
2	S ₀ -1	125	0.4	0.96
3	S ₀ -2	100	0.42	0.85
4	S ₀ +1	175	0.33	1.04
5	S ₀ +2	200	0.31	1.06
6	S ₀ -1	125	0.4	0.96
7	S ₀	150	0.35	1
8	S ₀ +1	175	0.33	1.04
9	S ₀ +2	200	0.31	1.06
10	S ₀ -2+ΔS	111	0.41	0.92
11	S ₀ -2	100	0.42	0.85
12	S ₀ -2+ΔS	111	0.41	0.92
13	S ₀	100	0.42	0.85
14	S ₀ -2-ΔS	89	0.42	0.83

Elaborazione risultati sperimentali isotensione

Come detto, il metodo JSME fornisce la procedura per l'elaborazione statistica dei risultati ottenuti dalle prove di fatica. La procedura proposta consente anche di calcolare le due rette della curva di Wohler in scala semilogaritmica con circa lo stesso livello di confidenza, pari al 50% di sopravvivenza.

Per il tratto inclinato della curva, note le coppie tensione applicata – numero di cicli il metodo consente di ricavare i parametri della retta interpolante:

$$\log(N) = \hat{\alpha}_1 + \hat{\beta}_1 \sigma \quad (1)$$

dove N è il numero di cicli e σ è la tensione applicata al filo SMA. Inoltre è possibile ricavare anche la stima della deviazione standard della vita a fatica logaritmica, $\text{dev}(\log N)$, e della tensione, $\text{dev}(\sigma)$.

Per il tratto orizzontale della curva di Wohler il metodo fornisce il livello di tensione limite di fatica a 5×10^5 cicli come la media dei livelli di tensione delle prove staircase. La confidenza del dato che si ricava è al 50%.

3. RISULTATI

In Tab. 4 sono riassunti i risultati sintetici delle prove di fatica in isotensione, per ognuno dei 14 livelli di carico sperimentati. La tabella, oltre a riportare il numero di cicli compiuti dal filo a rottura o al momento della sospensione della prova, riporta anche le principali grandezze meccaniche (variazione tensione meccanica, deformazione austenitica e martensitica, recupero) relative al ciclo 5 e al ciclo precedente il termine della prova.

In Fig. 6 si riporta la curva di Wohler ricavata secondo il metodo JSME (paragrafo Pianificazione prove sperimentali isotensione) sovrapposta ai punti sperimentali. Sul diagramma sono inoltre indicate le equazioni matematiche $\sigma - N$ dei due tratti.

Tabella 4. Risultati sintetici ottenuti dalle prove in isotensione

Prova	σ (MPa)	N_f	Ciclo 5				Ciclo (rottura/fine prova – 1)			
			$\Delta\sigma$ (%)	ε_{min} (%)	ε_{max} (%)	SME (%)	$\Delta\sigma$ (%)	ε_{min} (%)	ε_{max} (%)	SME (%)
1	150	9,688	3.8	0.50	4.93	4.43	4.4	1.01	5.39	4.38
2	125	25,901	4.4	0.47	4.64	4.17	4.5	0.72	4.84	4.12
3	100	606,214*	2.7	0.50	4.51	4.01	3.2	0.95	4.61	3.65
4	175	5,783	3.6	0.60	5.15	4.55	4.4	1.48	5.96	4.49
5	200	4,940	4.5	0.79	5.47	4.68	4.0	2.34	6.92	4.58
6	125	22,449	4.3	0.71	4.94	4.23	4.0	0.97	5.15	4.18
7	150	7,731	3.7	0.56	5.03	4.47	4.5	1.01	5.39	4.39
8	175	5,354	5.7	0.81	5.33	4.52	4.3	1.70	6.17	4.47
9	200	3,509	5.4	0.68	5.29	4.61	3.8	2.59	7.06	4.47
10	111	51,802	3.8	0.41	4.47	4.06	4.0	0.63	4.61	3.98
11	100	507,873*	4.6	0.60	4.54	3.94	3.0	1.02	4.64	3.62
12	111	32,923	4.1	0.36	4.49	4.13	4.3	0.53	4.50	3.97
13	100	389,614	3.2	0.47	4.40	3.93	3.4	0.84	4.45	3.61
14	89	507,176*	5.7	0.43	4.41	3.98	2.5	0.74	4.39	3.65

*Prova terminata senza rottura del filo

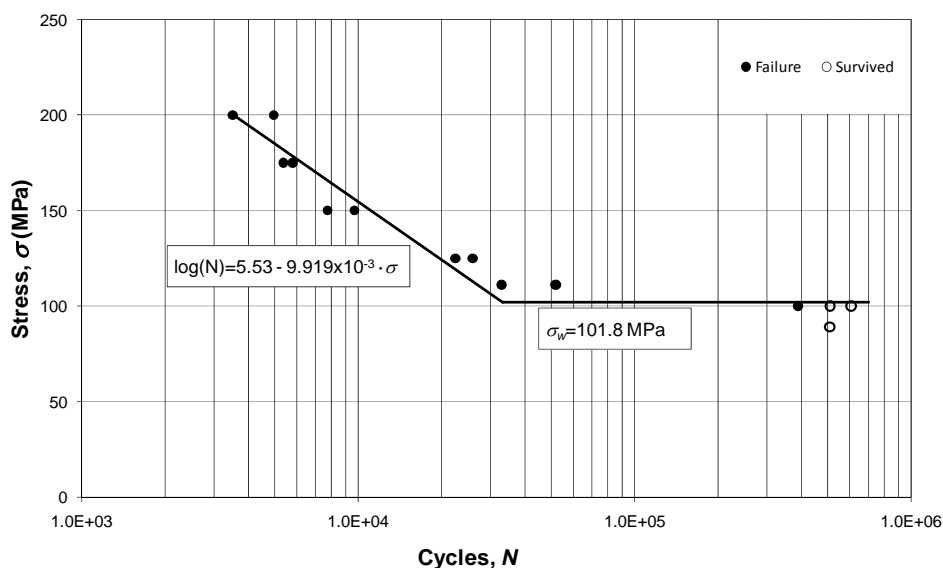


Figura 6. Curva di Wohler ricavata secondo il metodo JSME, completa delle equazioni individuate per i due tratti inclinato e orizzontale

La relazione tra tensione applicata al filo e deformazione (austenitica, martensitica, recupero) è visibile in Fig. 7a e 7b, relative rispettivamente al ciclo 5 e al ciclo precedente il termine della prova.

Per alcuni livelli di tensione (100MPa, 125MPa, 150MPa, 200MPa) si riporta in Fig. 8 la variazione delle deformazioni in funzione del numeri di cicli effettuati dal filo.

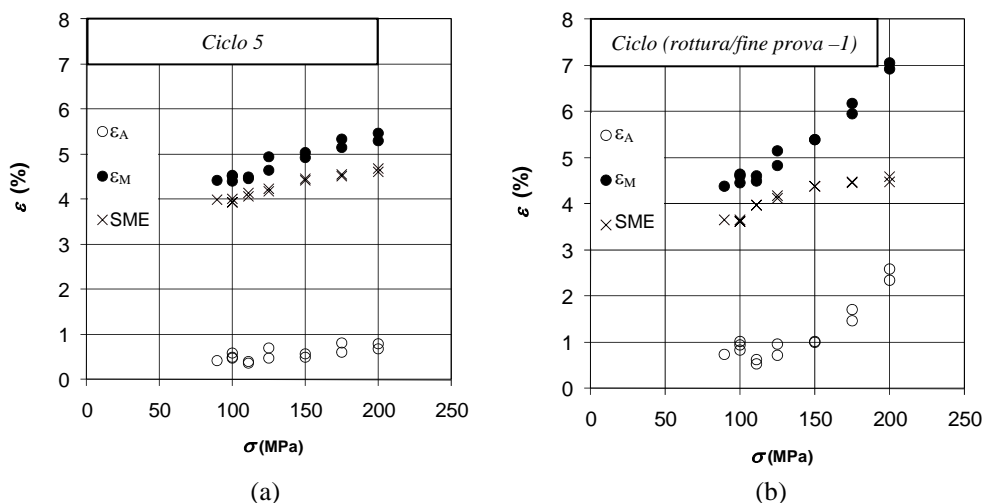


Figura 7. Deformazione in stato austenitico, martensitico e recupero in funzione della tensione di prova: a) al ciclo 5, b) al penultimo ciclo di prova

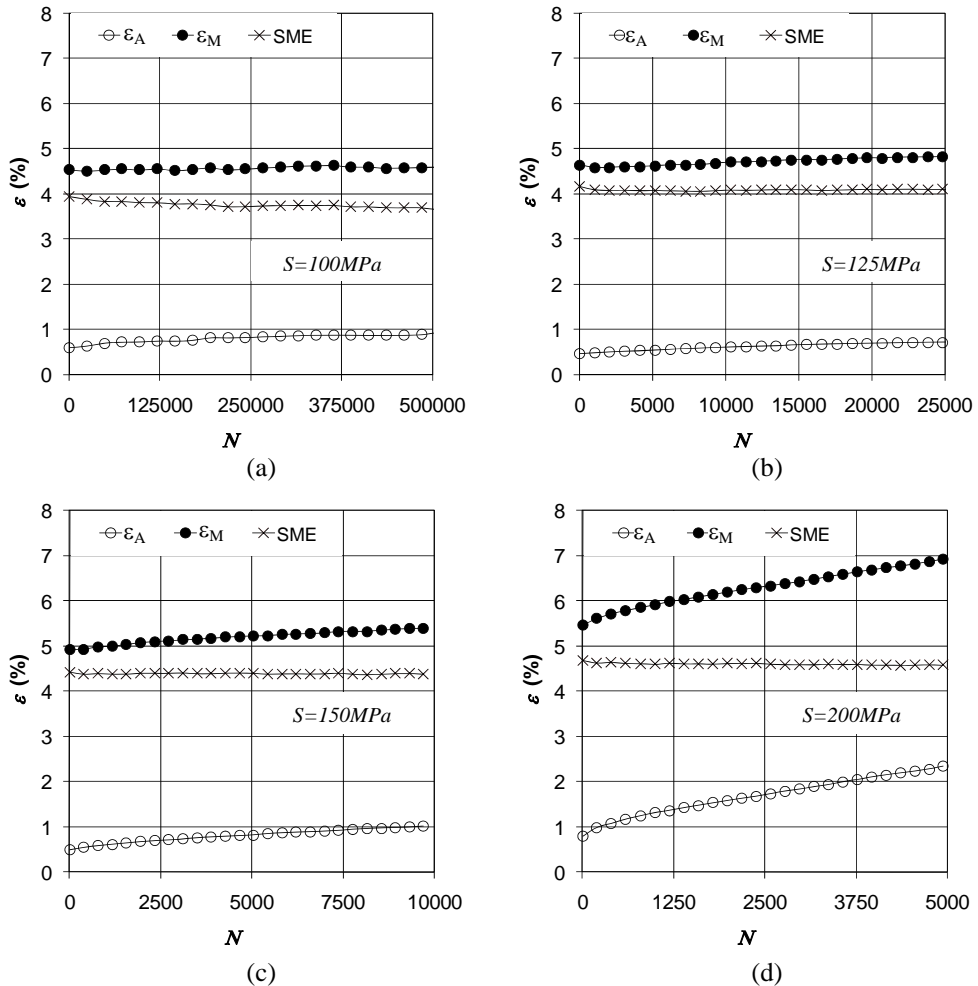


Figura 8. Evoluzione della deformazione in stato austenitico, martensitico e del recupero di deformazione per differenti valori di tensione applicata: a) 100MPa, b) 125MPa, c) 150MPa, d) 200MPa

Infine, in Fig. 9 è graficata la deriva di lunghezza del filo, intesa come la differenza fra la deformazione martensitica al ciclo N e la stessa deformazione registrata al ciclo 5. Questa grandezza, seppur non sia propriamente la deformazione plastica accumulata dal filo, ne è un surrogato ben rappresentativo.

4. DISCUSSIONE

Le prove in isotensione sono state condotte seguendo il metodo proposto dalla normativa giapponese [13], utilizzando la macchina di prova sviluppata. Per tutti i 14 livelli suggeriti dalla normativa si sono condotte prove preliminari atte all'individuazione dei parametri ottimali di alimentazione del filo, che garantissero la rapidità di esecuzione,

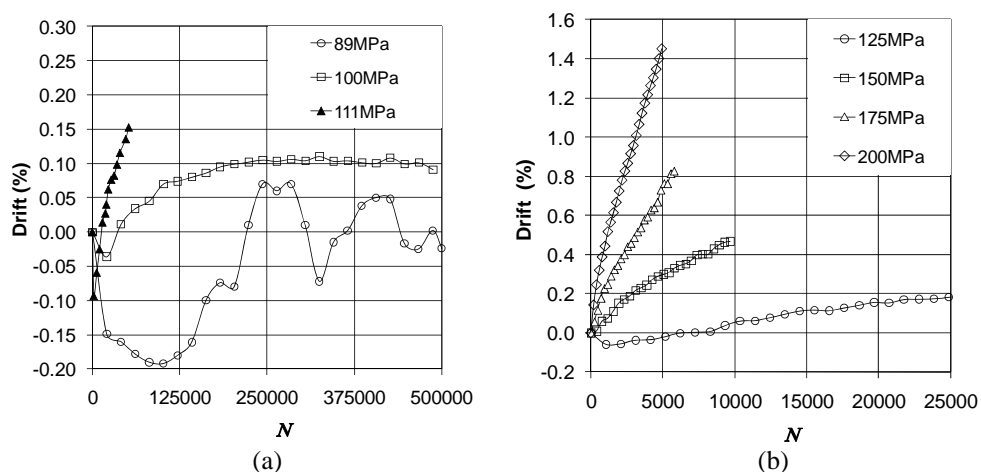


Figura 9. Evoluzione dell'allungamento del filo (Drift) durante le prove: a) bassi carichi, b) alti carichi

la completezza delle trasformazioni cristalline, impedissero il sovrariscaldamento del filo e garantissero la condizione di isotensione (Tab. 3). Da una verifica a fine prova dei dati acquisiti tutti questi obiettivi sono stati raggiunti, a garanzia dell'attendibilità dei risultati. In particolare, è stato raggiunto per tutti i cicli di tutte le prove l'obiettivo di limitare la variazione di tensione meccanica $\Delta\sigma$ nel ciclo a circa il 5% della tensione nominale (Tab 4).

Tutte le prove hanno fatto registrare rotture del filo nella zona centrale del filo o comunque ad una distanza mai inferiore ai 10mm dagli afferraggi.

Come si può osservare in Tab. 4 il valore più alto di tensione al quale non si è registrato la rottura del filo è stato 100MPa. A questo livello si è registrata anche una rottura del filo a circa 389.000 e una non rottura fino a 606.000.

Da un'analisi della curva di Wohler e dei punti sperimentali (Fig. 6) si può osservare come la curva presenta un tratto inclinato molto pendente, a testimonianza che aumentando la tensione applicata la durata del filo decresce velocemente. Si nota inoltre dallo scostamento dalla retta inclinata dei punti a 125MPa che l'interpolazione lineare non è ottimale. L'equazione della retta interpolante mostra una pendenza $\sigma/\ln(N)$ di circa 101 MPa (con deviazione standard 11MPa), prossima alla pendenza di 117 MPa individuata da Araujo et al. [9] per una lega Ti.45.0Ni-5Cu(%at). Il limite di fatica individuato vale circa 101.8MPa. Tale valore è riferito ad un criterio di superamento prova di 500.000 cicli e ha una confidenza del 50%. Osservando i punti sperimentali del grafico di Fig. 6 si nota bene come l'ipotesi che il materiale presenti un limite di fatica è più che probabile. Per avallare questa tesi si pensa in futuro di riattivare le prove in cui il filo non ha subito rotture. Se si esaminano le durate ottenute con quelle dichiarate dal produttore del filo (Tab. 1) emerge una discrepanza considerevole. I dati disponibili, relativi ad una tensione applicata 170MPa (Tab. 1), forniscono una durata maggiore di 100.000 cicli e una deriva di 0.17%, contro i circa 10.000 cicli (Tab. 4) e lo 0.45% di deriva (Fig. 9b) ottenuti nella prova a 150MPa. Escludendo che vi possano essere differenze di prestazioni così marcate fra lotti di produzione differenti l'unica spiegazione può risiedere nella differente modalità di prova. I

dati forniti da SAES Getters® si riferiscono ad una modalità di prova che non prevede la trasformazione completa del filo, bensì una trasformazione martensite – austenite parziale, fino ad arrivare ad corsa attuata del 3.5%. Le prove da noi effettuate sono invece relative ad una trasformazione completa, che nel caso del livello 150MPa porta ad una corsa di circa il 4.4%. Per fili in condizioni di isotensione, il beneficio della trasformazione parziale sulla durata del filo è già stato studiato e confermato da altri autori [4],[6]. Va sottolineato però che in questi lavori la trasformazione parziale è ottenuta facendo lavorare il materiale sempre all'interno della zona di isteresi, senza arrivare mai ad una completa trasformazione di fase. Sarebbe molto interessante studiare l'influenza, oltre che del grado di trasformazione, anche della posizione della trasformazione all'interno della zona di isteresi.

In Fig 7a sono visibili le deformazioni in stato austenitico, martensitico e la deformazione recuperata al ciclo 5. Si osserva come tutti i punti relativi alla deformazioni in stato austenitico ε_A siano interpolabili ottimamente con una retta, circostanza spiegabile per il fatto che il materiale in questo stato ha un comportamento pressoché perfettamente elastico (Fig. 5a). Anche i punti relativi alla deformazioni in stato martensitico ε_M sono ben interpolabili con una retta, conseguenza del fatto che nell'intervallo fra 100 e 200 MPa il materiale ha un comportamento abbastanza lineare (Fig. 5b). Ne consegue che anche la deformazione recuperata SME aumenti in modo lineare con la tensione applicata anche se risulta evidente che tra 175Mpa e 200Mpa l'incremento del recupero sia molto esiguo. Questo potrebbe essere spiegabile con il fenomeno riscontrato da Bignon et al. in [8], ossia che il valore massimo dell'effetto memoria di forma per una lega NiTi si raggiunge per una tensione di circa 200Mpa.

Con l'accumularsi dei cicli, il comportamento della lega cambia notevolmente. Facendo riferimento a Fig. 7b, relativa al penultimo ciclo di prova, si osserva un calo sensibile della deformazione recuperata SME per le prove a basso carico (fino al livello 110MPa), mentre per quelle ad alto carico (da 125MPa) la diminuzione della prestazione è più contenuta. Per contro, ad alti carichi il filo subisce allungamenti e quindi deformazioni decisamente importanti (Fig. 9b) dovuti all'accumulo di deformazioni plastiche permanenti. Per i punti a basso carico invece l'aumento della deformazione in stato martensitico è modesta e conseguentemente l'accumulo di deformazione permanenti è limitato (Fig. 9a). Ciò significa che la diminuzione della deformazione recuperabile prima osservata è imputabile ad una perdita di efficienza dell'effetto memoria di forma. Da un punto di vista puramente funzionale, è come se il modulo elastico austenitico diminuisse con il passare dei cicli. Questa osservazione è corroborata anche dal tracciato di Fig. 8a, relativo all'andamento delle tensioni in funzione del numero di cicli effettuati per un carico applicato di 100MPa: la deformazione martensitica è infatti pressoché costante, mentre la deformazione austenitica cresce linearmente, facendo calare con il medesimo andamento la deformazione recuperata. A carichi più alti (Fig. 8b,c,d), l'andamento in funzione dei cicli effettuati della deformazione martensitica ed austenitica è molto più simile e la deformazione recuperata è pressoché costante.

La deriva del filo, intesa come la differenza fra la deformazione martensitica al ciclo N e la stessa deformazione registrata al ciclo 5, al variare del numero di cicli accumulati presenta un andamento abbastanza singolare, almeno a carichi medio-bassi. Si nota infatti (Fig. 9a) come ad inizio prova la deriva sia negativa, ossia la deformazione massima martensitica cala rispetto al ciclo 5 man mano che il filo accumula cicli. Questo potrebbe essere spiegabile come un assestamento del filo dopo i trattamenti termo meccanici subiti

nelle fasi training. Nel caso di alti carichi invece (Fig. 9b), questo fenomeno non si evidenzia.

5. CONCLUSIONI

Il lavoro illustra una nuova metodologia di prove per la caratterizzazione a fatica di fili a memoria di forma, con l'obiettivo di ricavare dati più aderenti alle reali condizioni di funzionamento dei fili negli attuatori.

Il metodo prevede di effettuare prove di fatica secondo quattro differenti modalità di prova: isotensione, isodeformazione, isotensione con deformazione limitata, ciclo di tensione con deformazione limitata.

Una campagna sperimentale in linea con il metodo, compiuta su un filo commerciale, è attualmente in corso. I risultati fino ad ora ottenuti, ricavati con la macchina di prova appositamente realizzata, sono relativi alla condizione di isotensione.

Dall'analisi statistica dei risultati, eseguita secondo il metodo giapponese JSME, emerge come la curva di Wohler del materiale abbia un tratto inclinato a pendenza elevata. Dalle analisi condotte considerando un obiettivo di fine prova di 500.000 cicli, il materiale sembra presentare un limite di fatica di circa 100MPa.

Rispetto ai dati forniti dal produttore del filo, i risultati ottenuti sono in generale meno ottimistici. La discrepanza è quasi sicuramente imputabile al fatto che le prove qui effettuate sono relative ad una trasformazione cristallina completa, mentre nel caso del produttore le prove sono state eseguite con una trasformazione martensite-austenite parziale.

Attualmente sono in corso di svolgimento le prove in isodeformazione.

BIBLIOGRAFIA

- [1] Spinella I, Dragoni E., 2009. "Design equations for binary shape memory actuators under dissipative forces", *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Scienc*, **223**, pp 531–543.
- [2] Fumagalli L, Butera F., Coda A, 2009. "SmartFlex NiTi Wires for Shape Memory Actuators", *J. of Materials Engineering and Performance*, **18**, pp 691–695.
- [3] Mertmann M, Vergani G., 2008. "Design and application of shape memory actuators", *Eur. Phys. J. Special Topics*, **158**, pp 221–230.
- [4] Bertacchini O. W., Lagoudas D.C., Patoor E., 2009. "Thermomechanical transformation fatigue of TiNiCu SMA actuators under a corrosive environment – Part I: Experimental results", *International Journal of Fatigue*, **31**, pp 1571–1578.
- [5] G. Eggeler G., Hornbogen E., Yawny A., Heckmann A., Wagner M., 2004. "Structural and functional fatigue of NiTi shape memory alloys", *Materials Science and Engineering A*, **378**, pp 24–33.
- [6] Lagoudas D.C., Miller D. A., Rong L., Kumar P.K., 2009. "Thermomechanical fatigue of shape memory alloys", *Smart Materials and Structures*, **18**, 085021.
- [7] Bignon M.J., Morin M., 1995. "Fatigue of the shape memory effect in thin wires-comparison between TiNi and CuZnAl", *Journal de physique IV*, **C2**, pp 385-390.
- [8] Bignon M.J., Morin M., 1996. "Thermomechanical study of the stress assisted two way memory effect fatigue in TiNi and CuZnAl wires", *Scripta materialia*, **35**, pp 1373-1378.

- [9] Araujo C.J., Morin M., Guerin G., 1997. "Fatigue behavior of Ti-Ni-Cu thin wires SME", *Journal de physique IV France*, **C5**, pp 501-506.
- [10] Mertmann M., Bracke A. Hornbogen E., 1995. "Influence of the thermal treatment on the stability of partially constrained recovery of NiTi Actuator wire", *Journal de physique IV*, **C8**, pp 1259-1264.
- [11] Demers V., Brailovski V., Prokoshkinb S.D., Inaekyana K.E., 2009. "Thermomechanical fatigue of nanostructured Ti–Ni shape memory alloys", *Materials Science and Engineering A*, **513-514**, pp 185-196.
- [12] Zanotti C., Giuliani P., Tuissi A., Arnaboldi S., Casati R., 2009. "Response of NiTi SMA wire electrically heates", *Esomat 2009*, **06037**, DOI:10.1051/esomat/200906037.
- [13] Nakazawa h, Kodama S., 1990. "Statistical S-N testing method with 14 specimens: JSME standard method for determination of S-N curves", *Statistical research on fatigue and fracture*, pp 59–69.

EFFECT OF BONDLINE THICKNESS ON THE STATIC STRENGTH OF STRUCTURAL ADHESIVES UNDER NEARLY-HOMOGENEOUS SHEAR STRESSES

Davide Castagnetti

*Department of Engineering Sciences and Methods,
University of Modena and Reggio Emilia, Italy
E-mail: davide.castagnetti@unimore.it*

Andrea Spaggiari

*Department of Engineering Sciences and Methods,
University of Modena and Reggio Emilia, Italy
E-mail: andrea.spaggiari@unimore.it*

Eugenio Dragoni

*Department of Engineering Sciences and Methods,
University of Modena and Reggio Emilia, Italy
E-mail: eugenio.dragoni@unimore.it*

Abstract. A usual experimental observation retrieved in the technical literature is that the strength of an adhesive joint decreases by increasing the adhesive layer thickness. This well known behaviour is still not completely understood. All the surveys found in literature consider a complex stress state in the adhesive with mode mixing, stress concentrations on the midplane and stress singularities at the interface occurring at the same time. This paper aims at estimating the effect of the adhesive thickness on its intrinsic static shear strength and evaluate if this strength can explain the behaviour of a real bonded joint. A nearly-uniform shear stress distribution is obtained through an ad-hoc tubular butt joint subject to pure torsion. A standard single lap joint is considered as benchmark, due to its complex and singular stress field into the adhesive. The experimental campaign is focused on two adhesives: a modified methacrylate and high strength epoxy. Four levels and three levels of the adhesive thickness were considered in the tubular butt joint and in the single lap joint respectively, all in the range between 0.05 and 0.4 mm. The effect of the adhesive thickness on the static strength of the adhesive is investigated by considering the type of failure and by comparing the structural stresses in the tubular butt joint with the ones on the midplane of the adhesive layer in the single lap joint. Moreover, the stress intensity factor in the single lap joint is calculated.

Keywords: structural adhesive, bondline thickness, static strength, shear stresses

1. INTRODUCTION

This paper deals with the analysis of stresses in quasi-static loaded bonded joints. In particular, the aim of the work is to assess the effect of the adhesive thickness on the static shear strength of the adhesive under a nearly-uniform stress distribution. In addition, the work investigates if this intrinsic shear strength can explain the decrease in the strength experimentally observed in a real bonded joint when the adhesive thickness increases. The technical literature reports contrasting theories about the relation between the theoretical adhesive stress distribution and the experimental strength. Dolev et al [1], on the basis of experimental tests on Napkin Ring specimens, show a modest effect of the adhesive layer thickness on the static shear strength. Crocombe [2] proves that, in single lap joints (SLJs), the strength decreases as the adhesive thickness increases, in particular when plasticity of the adhesive is considered. On the contrary, elastic analyses testify that a more uniform stress distribution is obtained along the bondline with increasing adhesive thickness [2]. Still, for SLJs, computational analyses provide higher interface stresses as the adhesive thickness increases, as found by Gleich et al [3]. Thus, a failure criterion relying on interface stresses is suggested in case of high adhesive thickness. The experimental tests on SLJs reported by Da Silva et al [4] testify increasing shear strength as the adhesive thickness decreases. Goglio et al. [5-7] show that the interface stresses increase as the adhesive layer thickness increases, so that the strength decreases. In a former theory, Adams and Peppiant [8] ascribe the smaller strength of the thicker adhesive to the higher probability of defects such as voids or micro-fractures inside the bondline. Adnan and Sun [9], dealing with joints subject to normal or shear stresses, impute to a better chemical and molecular behaviour the significant increase of the failure strength as the adhesive thickness decreases.

Hence, from the examined literature, it is possible to perceive both an elasto-mechanical and a chemical explanation of the decreasing strength as the adhesive thickness increases. The first links a thicker adhesive layer with higher interface stresses and therefore to a lower strength of the joint [2-7]. The second correlates a thinner adhesive layer to a better polymerization of the adhesive itself and therefore to stronger molecular bonds [4, 9].

A common limitation of the cited literature is to consider a complex stress state in the adhesive with mode mixing, stress concentrations on the midplane and stress singularities at the interface occurring at the same time. This work aims to overcome these drawbacks by measuring the adhesive shear strength as a function of the thickness, under a nearly-regular stress distribution. The nearly-uniform and singularity free shear stress distribution is obtained exploiting ad-hoc tubular butt joint (TBJ) geometry subject to torsional loading [10-11]. A systematic experimental campaign compares the static strength in the TBJ with the one in a SLJ assumed as benchmark configuration. Two structural adhesives (Loctite Multibond 330 [12] and Loctite Hysol 9514 [13]) are examined by considering four levels of the adhesive thickness in the TBJs and three levels, ranging in the same interval, in the SLJs. The TBJ is loaded in pure torsion while the SLJ is loaded under axial load, until failure occurs. Multibond 330 adhesive exhibited a typical cohesive failure in both the joints. By contrast, Hysol 9514 failed close to the interface.

The torsional tests on the TBJ show a significant effect of the adhesive thickness on the static strength of Multibond 330, while the Hysol 9514 adhesive exhibits a nearly constant static strength. With regard to the SLJ, both the structural stresses on the mid-plane of the adhesive and the stress intensity factor (SIF) linked to the stress concentrations on the

adherend-adhesive interface were examined. The results show that in the case of Multibond 330 the intrinsic shear strength in the TBJ can explain the decreasing strength in the SLJ as the adhesive thickness increases. By contrast, in the case of Hysol 9514 only the mean shear stress in the SLJ shows a good agreement with the one in the TBJ, while the structural stresses exhibit a noticeable discrepancy. The SIF in the SLJ has an opposite trend with respect to the experimental strength, thus it does not support the observation of a decreasing strength as the adhesive thickness increases.

2. MATERIALS AND METHODS

The work is organized in two steps. The first step describes the design of an ad-hoc joint specimen. The second step deals with the design, execution and analysis of the experimental tests.

Joint Design

The design of the tubular joint pursues two different aims. First, the TBJ originates a nearly uniform and singularity-free shear stress field in the adhesive, in order to clearly assess the effect of the adhesive thickness on its static strength. Second, the TBJ ensures the execution of different levels of the adhesive thickness, with both accuracy and simplicity. This second aim deals with the need of a simple to manufacture and inexpensive specimen, which allows the adhesive thickness to be easily set up.

Figure 1a displays a three-dimensional exploded view of the proposed joint, which is made up by two tubular adherends (*A* and *B*) having a different wall thickness. A Teflon (PTFE) insert (*T*) is used to connect the adherends. The thickness of the bonded region (Fig. 1b) is controlled by the length of the top of the insert (*T*) (Fig. 1b). Moreover, the Teflon insert ensures that the adherends are respectively co-axial and prevents the adhesive to flow into the joint. The insert (*T*) is made of Teflon due to its remarkable anti-adherent properties, in order to guarantee that adhesion occur only at the upper and lower adherend-adhesive interface. To assess the anti-adherent properties of Teflon with the adhesives here used, preliminary experimental tests were performed on TBJs completely made of Teflon. The results, not reported here for the sake of brevity, have confirmed this hypothesis. Therefore, any type of perturbation of the stress distribution in the adhesive layer is prevented. The transversal holes in the adherends are needed to fix the joint to the testing machine. A peculiarity of the proposed joint is the good accuracy of the adhesive layer thickness, achieved by minimizing the chain of tolerances involved in its manufacture (Fig. 1b). The desired adhesive thickness is obtained simply by combining two dimensions: the height of the top of the insert (*T*), and the depth of the hole in adherend *A* which is constant for all the joints. In particular, the dimension *X* in the Teflon insert (Fig. 1b) is the nominal value of the adhesive thickness.

The uniformity of the stress distribution in the adhesive layer for a completely symmetric TBJ configuration has been proved by different authors [10-11]. Since the proposed configuration has a different wall thickness between the adherends and the torque load is applied through the holes for the fixture, uniformity of the stress state is not assured. Therefore a FE analysis was performed which reproduced in details the three-dimensional geometry of the TBJ (Fig. 1a) with a thickness of the adhesive layer equal to 0.2 mm. Two layers of linear, hexahedral, solid elements with a cubic shape described the adhesive layer. The same element type was adopted for the adherends.

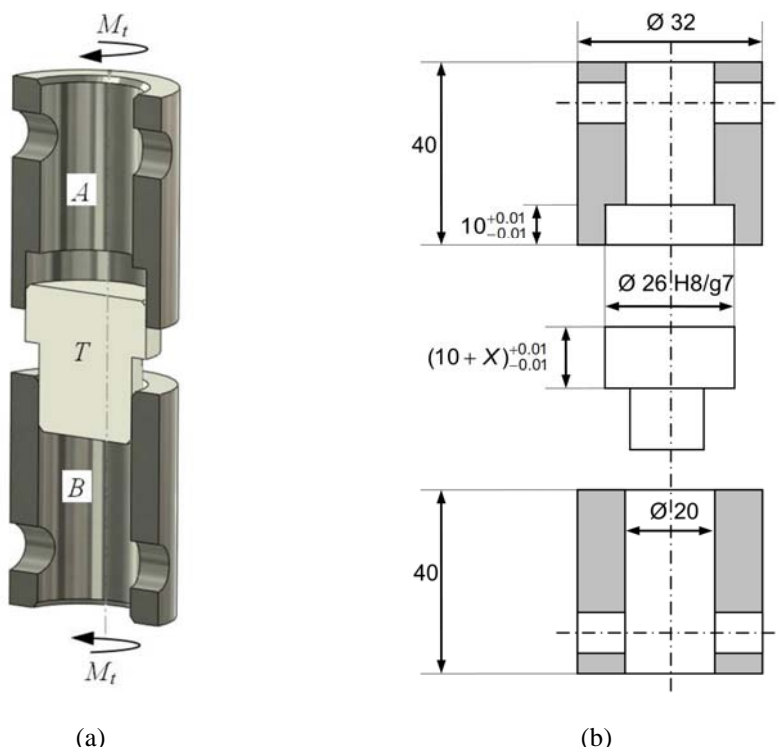


Figure 1. Exploded view (a) and technical drawing (b) of the proposed TBJ.

In particular, the same dimension of the elements as in the adhesive was chosen near the adhesive interface. Moving far from the adhesive, the element dimension was gradually increased and tetrahedral elements were adopted. A mild steel was assumed for the adherends having a Young's modulus of 210000 MPa and a Poisson's ratio of 0.3. A Young's modulus of 1700 MPa, and a Poisson ratio of 0.4 were assumed for the adhesive. These represent the typical elastic properties of structural adhesive.

Both the constraints and the torque load were applied to the transverse holes in the adherends, in order to resemble the actual test conditions. On the one hand, the constraints prevented rigid rotations of the joint by fixing two opposite surfaces of the holes on the bottom adherends. On the other hand, the torque load was obtained by applying an equivalent pressure to opposite surfaces of the transverse hole in the top adherend. The analysis was performed in the elastic field. The model, implemented in the ABAQUS 6.8 software [14], was run on a machine equipped with a dual-processor (Xeon 3.2GHz) and 4GB of RAM.

Experimental Campaign

A systematic experimental campaign was performed which examined two structural adhesives in two different joints (Tab. 1). The two adhesives considered in this work were all supplied by Henkel (formerly Loctite).

Table 1 Factors and factor levels of the systematic experimental campaign.

Adhesive layer thickness (mm)		Adhesive	
		Multibond 330	Hysol 9514
Joint	TBJ	0.05 , 0.1 , 0.2 , 0.4	
	SLJ	0.05 , 0.15 , 0.4	

They are chemically different, the first is the Multibond 330 [12], a two-part modified methacrylate ester and the second is the Hysol 9514, a heat-curing, single-component epoxy [13]. The TBJ proposed in Section 2.1, and a SLJ conform to ASTM D1002-01 were adopted. The SLJ was chosen as benchmark due to its highly irregular stress field in the adhesive where both stress concentrations (on the mid-plane) and stress singularities (on the adherend-adhesive interface) arise.

On the one hand, in the TBJ, four different values of the adhesive thickness were examined, corresponding to the X dimension in Fig. 1b: 0.05, 0.1, 0.2, and 0.4 mm. The thickness of 0.05 is the result of the adherends roughness, while the other thickness values were enforced by the X dimension of the Teflon insert. On the other hand, in the SLJ, only three thickness values (0.05, 0.15 and 0.4 mm) were considered. With exception of the 0.05 mm thickness, the remaining values were enforced introducing three short calibrated wires in the center of the overlap region. These values were chosen in the typical range of application for structural adhesives. The adherends of the TBJ were made of quenched steel 39NiCrMo3 (EN 10083-3 [15]), while the thin sheet adherends for single-lap joints were made of standard low carbon, cold-rolled steel complying with ASTM A-109 [16]. For each of the joint configurations, five replicates were performed giving a total of 50 tests on TBJs and 30 tests on SLJs.

All the tubular adherends and Teflon inserts, needed for the experimental campaign, were manufactured by turning. The depth of the cavity in the adherends and the calibrated height of the Teflon inserts (X in Fig. 1b) were measured with an altimeter (Mitutoyo 570-302, resolution 0.01 mm). Then the optimal precision of the thickness of the adhesive layer was obtained by combining the parts in order to minimize the error on this chain of tolerances. The joints execution was organized as follows. The surface preparation of the adherends was performed manually with sandpaper (P100) following a crossed path, in order to increase the adhesion properties of the bonding surfaces. The surface roughness was measured with an electronic rugosimeter Hommelwerke T-500 on a representative sample of 2 adherends for each adhesive thickness for each type of joint. In order to remove all the micro particles produced by the surface preparation and also to degrease the surfaces before the application of the adhesive, the adherends were finally cleaned with Loctite 7063. The assembly of the joint was performed carefully in order to avoid air inclusions in the adhesive layer. The excess of adhesive, squeezed out on the external side, was removed in order to obtain a straight and smooth lateral surface, with no fillets. The joints bonded with Multibond 330 adhesive were cured for 5 days at room temperature, while the joints bonded with Hysol 9514 were cured for 90 minutes in the oven at 175°C. Both the batches of TBJs and also the SLJs were executed in a randomized order.

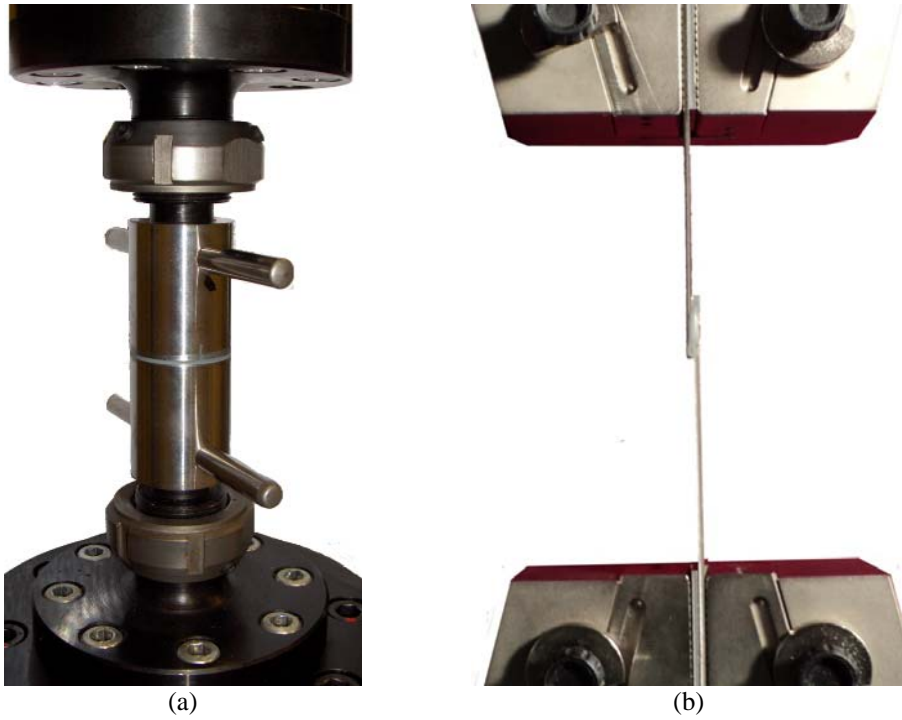


Figure 2. Set up of TBJs (a) and of SLJs (b) on the electro-hydraulic testing machine.

Quasi-static failure tests were performed by applying a torque load (0.1°/sec) to the TBJs and an axial load (1 mm/min) to the single-lap joints. The torque-angle curve and the traction-force curve up to failure were registered respectively. Figure 2a and b show the setup of the TBJ and of the SLJ respectively, on the electro-hydraulic testing machine (MTS MiniBionix 858) controlled by a PC equipped with proprietary data acquisition software.

3. RESULTS

Joint Design

The diagram in Fig. 3 shows, for the examined configuration of the joint (Section 2.1), a comparison between the computational (symbols) and analytical (solid line) shear stress in the adhesive layer as a function of the radius of the tubular adherend. The analytical stresses are calculated with the following formula:

$$\tau(r) = \frac{M_t \cdot r}{J_p} \quad (1)$$

where M_t represents the torque on the joint, r is the generic radius and J_p is the polar inertia moment of the annular section.

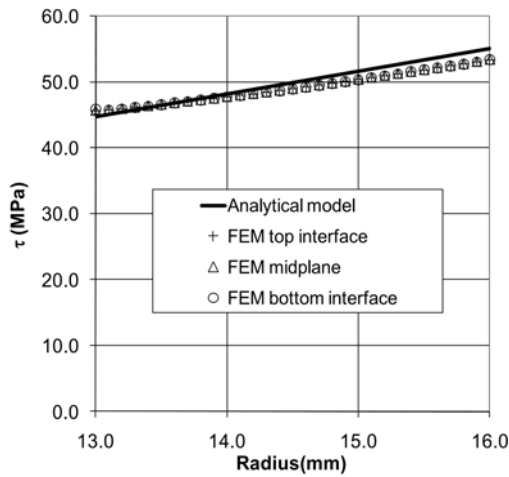


Figure 3. Analytical versus computational shear stress distribution along the radius of the annular section of the adhesive layer, in the TBJ.

The diagram provides the computational shear stresses both on the adherend-adhesive interfaces (crosses and empty circles on the top and bottom interface respectively) and on the mid-plane of the adhesive layer (empty triangles). These values correspond to an applied torque load equal to 200 Nm.

Experimental Campaign

Table 2 presents a summary of the adherends roughness in terms of the average roughness, R_a , of the average of the maximum roughness, R_z , and of the maximum roughness, R_{max} , accordingly to UNI ISO 4287. The standard deviation is provided for each value. Table 3 displays the adhesive layer thickness and the corresponding standard deviation and relative error, between the actual and nominal value, for each joint. The values were measured as described in Section 2.2.

Table 2 - Surface roughness of the bonding surfaces of the adherends after grit blasting.

	Ra (μm)	Rz (μm)	Rmax (μm)
Average roughness	0.72	4.41	5.89
Standard deviation	0.25	1.38	1.78

Table 3 - Measured adhesive layer thickness in the TBJs.

Nominal adhesive thickness	0.05	0.1	0.2	0.4
Average measured thickness (mm)	-	0.111	0.206	0.406
Standard deviation (mm)	-	0.00212	0.00152	0.00187
Relative error (%)	-	+11.0	+3.0	+1.5

Figure 4 shows a detail of the surface of the joints at failure. Figure 4a shows two joints bonded with Multibond 330 adhesive, having a thickness of the adhesive layer equal to 0.05 mm (on the bottom) and 0.4 mm (on the top) respectively. By keeping the same layout and adhesive layer thicknesses, Fig. 4b shows two joints bonded with Hysol 9514.

Figure 5 presents the diagrams of the experimental torque-angle curves for the TBJs bonded with Multibond 330, for all the four levels of the adhesive thickness. Five curves are presented for each diagram, corresponding to the replicates performed for each joint configuration. Similarly, Fig. 6 presents the same diagrams (starting from a preload torque of 20Nm) for the TBJs bonded with Hysol 9514.

Table 4 collects the peak torques retrieved from the experimental curves in the diagrams in Fig. 5 and 6 for all the TBJs tested in the experimental campaign. The points in Fig. 7 show the average value of the maximum shear stresses retrieved through equation (1) on the tubular adherends, calculated from the peak torques in Tab. 4. The empty circles indicate the average stress value for Multibond 330 adhesive, and the solid triangles represent the average value for the Hysol 9514 adhesive. The vertical lines in the diagram represent the confidence interval (plus or minus one standard deviation) of each average value.

Figure 8 presents the diagrams of the experimental force-displacement curves for the SLJs bonded with Multibond 330 and Hysol 9514 adhesive, for all the three levels of the adhesive thickness. For both the adhesives the curves starts from a preload of 1000N. Table 5 collects the experimental peak forces retrieved from the experimental curves in the diagrams in Fig. 8.

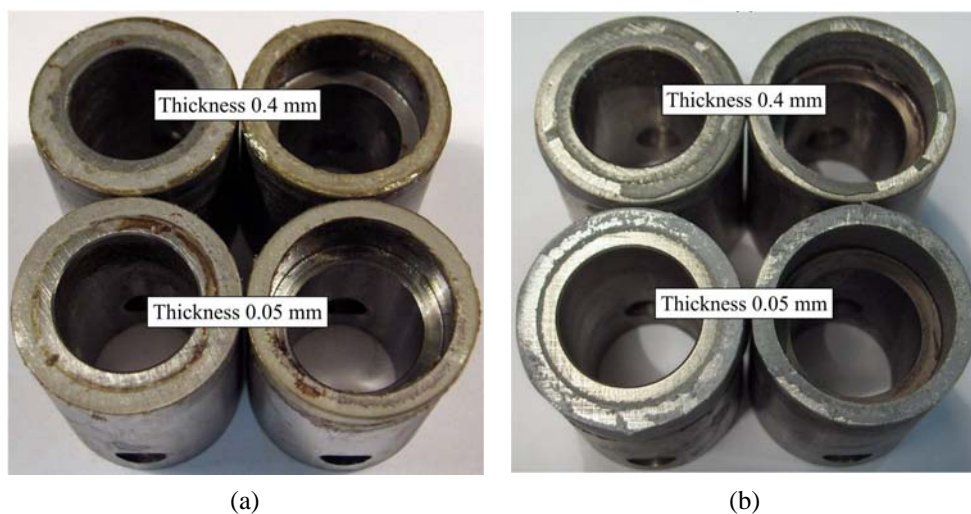


Figure 4. View of the failure surfaces of the adhesives for two joints bonded with Multibond 330 (a) and for two joints bonded with Hysol 9514 (b), both of them for the same values of the adhesive thickness.

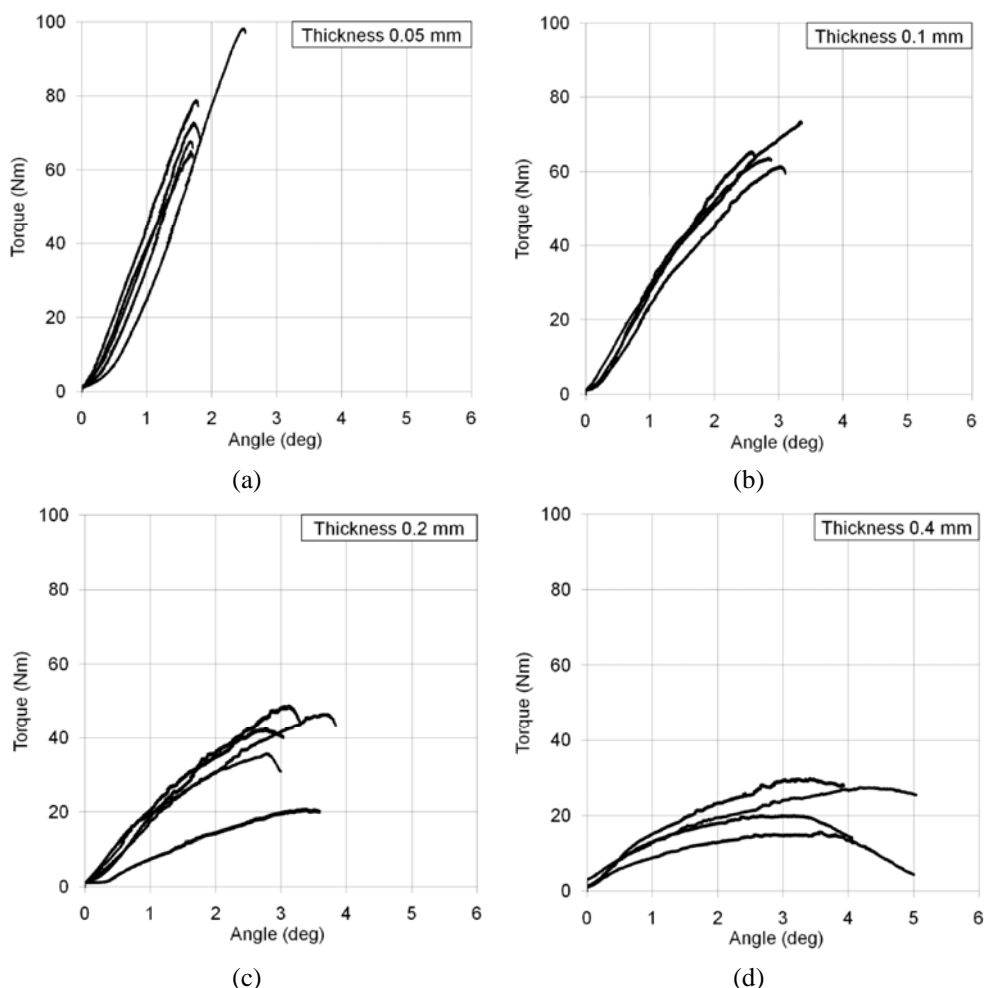


Figure 5. Torque-angle diagrams of the experimental tests on the TBJs bonded with the Multibond 330 adhesive, for the adhesive layer thickness 0.05mm (a), 0.1mm (b), 0.2mm (c) and 0.4mm (d).

The points in Fig. 9 display the average values of the peak equivalent stresses corresponding to the peak forces in Tab. 5, as a function of the adhesive layer thickness. The peak stresses were calculated (as explained in the following) on the midplane of the adhesive layer for all the SLJ configurations. The Rankine equivalent stress criterion was chosen for both the adhesive due to their brittle constitutive behaviour as documented by Goglio et al. [6] for Multibond 330, and by Peroni et al. [17] for Hysol 9514. The symbols here adopted are the same as in the previous diagrams. The vertical lines straddling each average value represent its confidence interval (plus or minus one standard deviation).

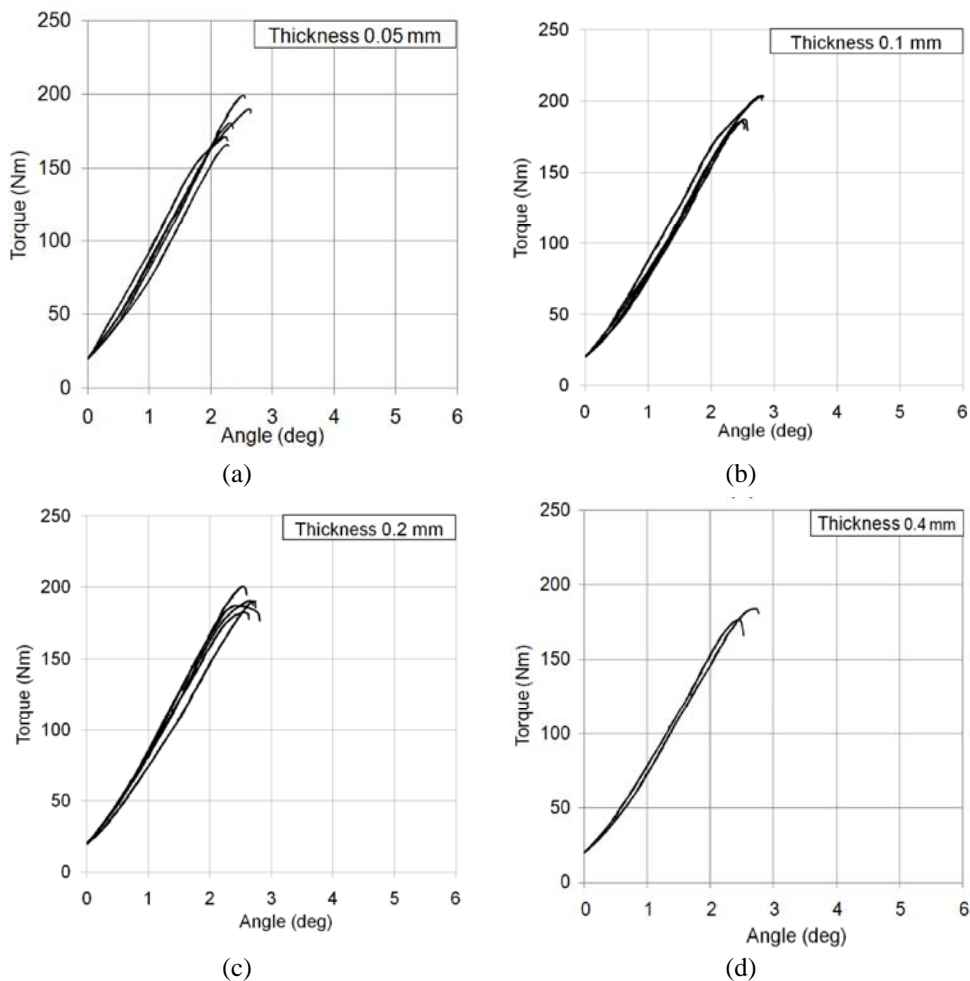


Figure 6. Torque-angle diagrams of the experimental tests on the TBJs bonded with the Hysol 9514 adhesive, for the adhesive layer thickness 0.05mm (a), 0.1mm (b), 0.2mm (c) and 0.4mm (d).

Due to the yielding of the adherends, which occurred in the experimental tests on SLJs, the Rankine equivalent stress was evaluated with a finite element analysis. An efficient finite element model was adopted to describe the examined joint configurations. This model conforms in all details to the technique proposed by the authors in [18] for the elastic analysis of thin walled bonded structures, and then extended to the failure analysis of T-peel joints [19-20]. Therefore, for the sake of brevity, the computational model is not described in the present paper.

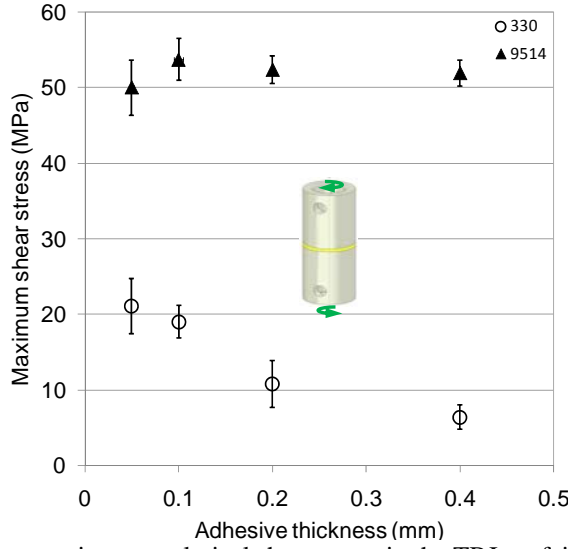


Figure 7. Average maximum analytical shear stress in the TBJs at failure, both for Multibond 330 (empty circles), and for Hysol 9514 (solid triangles) on the midplane of the adhesive layer.

Figure 10 compare the stresses in the SLJ with the ones in the TBJ as a function of the adhesive thickness, for the Multibond 330 adhesive. The diagram in Fig. 10a shows the mean shear stresses at failure: for the TBJ it was calculated according to Eqn. (1), by assuming r equal to the average radius of the annular section of the adhesive layer. For the SLJ the average shear stress was calculated as the ratio of the peak force to the overlap area of the joint. The diagram in Fig. 10b presents the maximum shear stresses at failure and the diagram in Fig. 10c displays the maximum Rankine stresses at failure. In the case of the TBJ the stress values in Fig. 11b and c were calculated analytically through Eqn. (1), while for the SLJ all these stress values were retrieved from the computational analyses. Similarly, Fig. 11 presents the same diagrams for the Hysol 9514 adhesive.

Table 6 collects the stress intensity factor (SIF) at the adherends-adhesive interface in the SLJs for both adhesives. This parameter was calculated as proposed by Wang et al. [21] according to the following expression:

$$K = \left[\sigma^* \cdot A(\nu) + \tau^* \cdot B(\nu) \right] t^\lambda \quad (2)$$

where σ^* and τ^* are the peel and shear stresses evaluated by means of the structural solution, t is the adhesive thickness, $A(\nu)$ and $B(\nu)$ are combination coefficients (functions of the adhesive Poisson's ratio only) given by the following equations:

$$\begin{aligned} A(\nu) &= 0.836 - 2.23\nu + 6.29\nu^2 - 9.64\nu^3 \\ B(\nu) &= 3.12 - 15.8\nu + 40.1\nu^2 - 37.6\nu^3 \end{aligned} \quad (3)$$

Table 4 - Experimental peak torques in the TBJs.

	Thickness (mm)	0.05	0.1	0.2	0.4
Multibond 330	Peak torque (Nm)	73.0	61.6	42.7	30.0
		98.3	80.6	21.0	15.8
		78.9	73.7	35.9	22.4
		65.2	63.9	46.7	20.2
		67.8	65.6	48.9	27.6
	Average peak torque (Nm)	76.6	69.1	39.1	23.2
	Standard deviation (Nm)	13.2	7.9	11.2	5.7
Hysol 9514	Peak torque (Nm)	189.7	194.7	186.9	172.1
		171.7	185.8	190.3	176.3
		180.8	203.7	182.4	161.9
		199.0	187.2	200.5	176.6
		166.1	203.8	190.0	184.1
	Average peak torque (Nm)	181.5	195.0	190.0	174.2
	Standard deviation (Nm)	13.3	8.6	6.7	8.1

Table 5 - Experimental peak forces in the SLJs.

	Thickness (mm)	0.05	0.15	0.4
Multibond 330	Peak force (N)	5395.8	6033.4	3498.8
		6104.2	5325.8	4350.0
		6036.2	5336.2	3993.5
		5648.8	5761.4	4549.7
		6579.5	5695.7	5353.0
	Average peak force (N)	5952.9	5630.5	4349.0
	Standard deviation (N)	454.2	301.3	688
Hysol 9514	Peak force (N)	12434.4	10247.9	9358.6
		12018.9	11677.5	10024.3
		10578.6	11775.7	9025.7
		11275.2	12536.6	9276.0
		10986.0	11328.2	9159.7
	Average peak force (N)	11458.6	11513.2	9368.9
	Standard deviation (N)	757.9	833.5	387.3

Finally, the singularity order λ has been calculated as the root of the following equation, given in [21] for a square-edged adhesive layer:

$$(3 - 4\nu)\cos(\pi(1 - \lambda)) - 2(1 - \lambda)^2 + 8\nu^2 - 12\nu + 5 \quad (4)$$

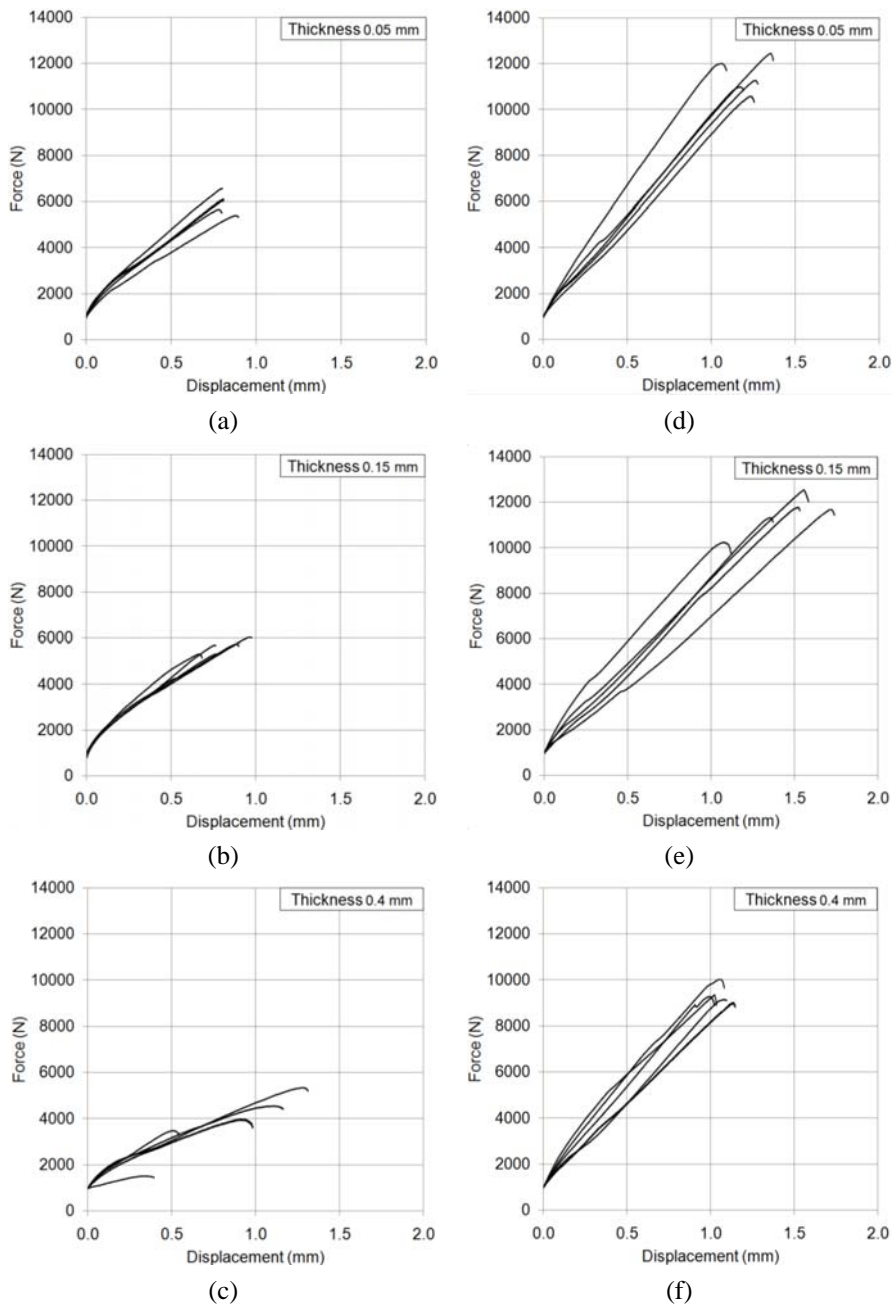


Figure 8. Force-displacement diagrams of the experimental tests on the SLJs bonded with the Multibond 330 adhesive.(a-b-c) and with Hysol 9514 (d-e-f) for the adhesive layer thickness 0.05mm, 0.15mm and 0.4mm.

Table 6 - Stress intensity factor K for the SLJs at failure.

Stress intensity factor, K	Thickness (mm)		
	0.05	0.15	0.4
Multibond 330	55.9	43.1	31.7
Hysol 9514	57.0	52.6	47.4

4. DISCUSSION

Joint Design

From the diagram in Fig. 3 it appears that the proposed TBJ, loaded by torque, originates a nearly-uniform shear stress distribution in the adhesive. In accordance with equation (1), solid line in the diagram, the shear stress depends linearly on the radius of the joint and is constant through the adhesive thickness. The same values originates both on the midplane of the adhesive layer (empty triangles) and on the adhesive-adherend interfaces (crosses and empty circles).

Experimental Campaign

As reported in Tab. 2 the average roughness of the bonding surfaces was quite constant and uniform between the specimens. The data about the thickness of the adhesive layer (Tab. 3) show a great accuracy and repeatability in the execution of the joints. The relative error between the actual and nominal thickness ranges from 11% for 0.1 mm up to a minimum of 1.5% for 0.4 mm.

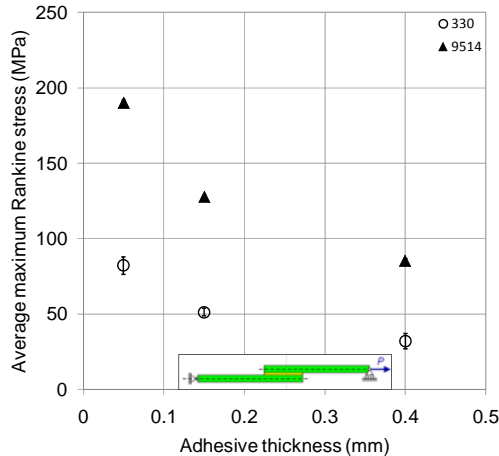


Figure 9. Average maximum Rankine equivalent stress in the SLJs at failure both for Multibond 330 (empty circles), and for Hysol 9514 (solid triangles) on the midplane of the adhesive layer.

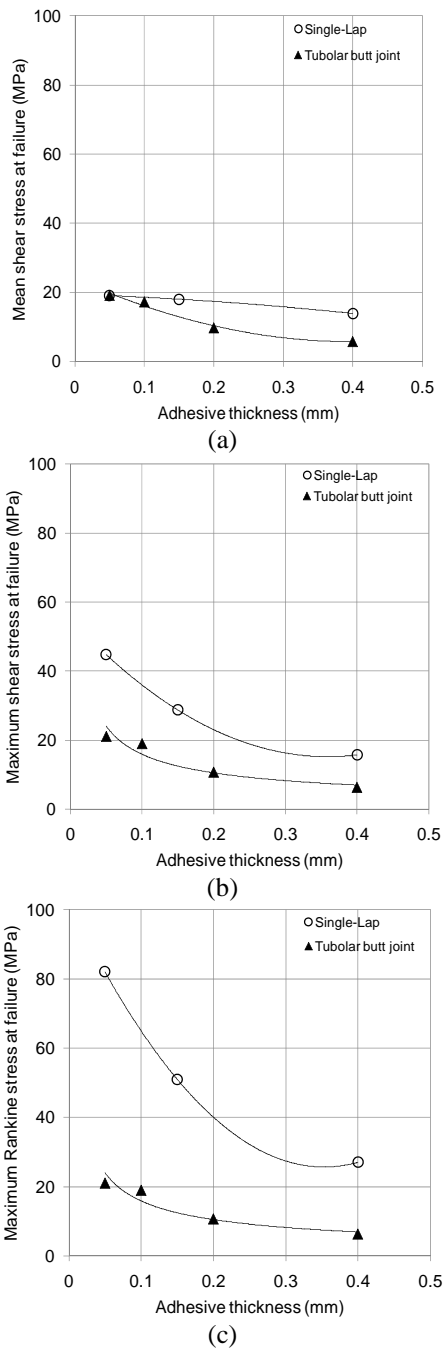
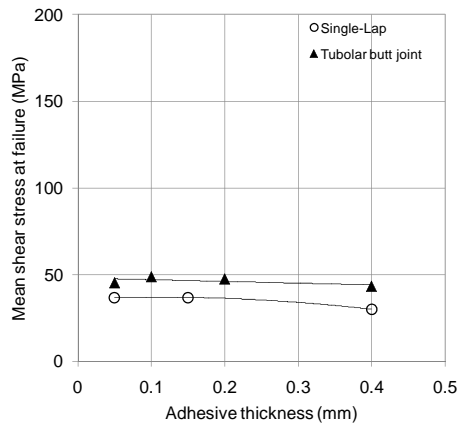
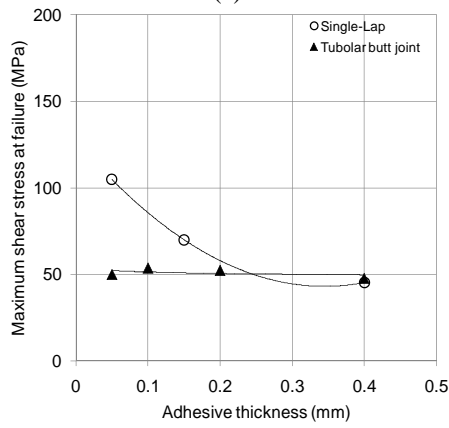


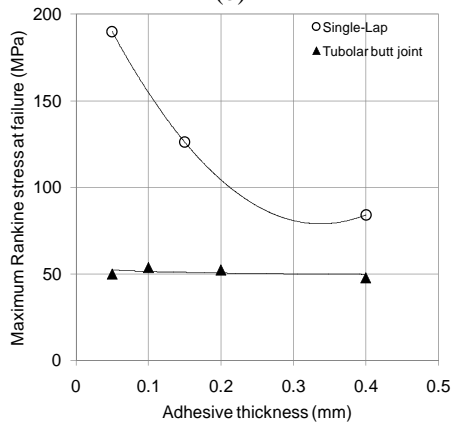
Figure 10. Comparison between the structural stresses in the TBJs and in the SLJs for Multibond 330: average shear stress (a), average maximum shear stress (b), and average maximum Rankine stress (c) on the midplane of the adhesive layer.



(a)



(b)



(c)

Figure 11. Comparison between the structural stresses in the TBJs and in the SLJs for Hysol 9514: average shear stress (a), average maximum shear stress (b), and average maximum Rankine stress (c) on the midplane of the adhesive layer.

By examining the surface of the TBJs at failure (Fig. 4) a different behaviour can be observed between the adhesives. Figure 4a shows a typical cohesive failure for the Multibond 330 adhesive, with a failure surface approximately on the midplane of the adhesive. In the case of the Hysol 9514 adhesive (Fig. 4b), failure occurs close to an interface and the failure surface crosses the adhesive layer many times, from one interface to the opposite, following a zigzag path. A quite similar failure behaviour of the adhesives, which is described also by Pirondi et al. in [22], was observed also in the SLJs, not reported here for the sake of brevity.

Figure 5 shows for the TBJs with Multibond 330 that the repeatability of the torque-angle curves slightly decreases as the thickness increases. Moreover, the curves highlight a remarkable decrease in the stiffness and strength of Multibond 330, as the adhesive thickness increases. The stiffness decrease may be attributed to the increase in the adhesive thickness. However, a more accurate measure of the strains occurring in the adhesive layer would be needed to assess this hypothesis. On the contrary, the strength decrease is probably imputable to lower mechanical properties of the adhesive itself. By contrast, Hysol 9514 (Fig. 6) shows a higher stiffness and peak torque, which are nearly constant in the examined range of thickness. Also in this case, a more accurate measure of the strains in the adhesive would be needed to investigate effect on the stiffness. With regard to the strength, this constant response is probably imputable to the almost-interfacial failure of these adhesive. The data reported in Tab. 4 point out that the peak torque of the Hysol 9514 is twice the maximum peak torque of Multibond 330 with a lower standard deviation.

By examining the static shear strength of the adhesives in Fig. 7 two observations can be drawn. On the one hand, Multibond 330 exhibits a remarkable decrease as the thickness increases: from a maximum of about 20 MPa for a thickness below 0.1 mm, the strength decreases down to about 6 MPa at 0.4 mm. On the other hand, Hysol 9514 provides nearly constant shear strength in the whole range of thicknesses, as can be deduced by the overlapping error bars. Since a nearly-uniform and singularity-free stress field occurs in the TBJ and cohesive failure is observed in Multibond 330, its decreasing strength by increasing thickness may be attributed to a lower mechanical property of the adhesive itself. By contrast, the nearly-constant strength observed for the Hysol 9514 can be attributed to its almost-interfacial failure, which causes breaking of the chemical bonding between the adherends and the adhesive. Therefore, the thickness cannot play a significant role.

The experimental curves in Fig. 8 show a lower repeatability of the SLJs in comparison with TBJs. Similarly to the TBJs, Multibond 330 provides, in the SLJs, stiffness about one half the one of Hysol 9514 and lower peak loads. Table 5 points out that the average peak forces for Multibond 330 are about one half the ones for Hysol 9514 and a quite similar standard deviation is obtained. On the whole, these results are in accordance with the data provided by the manufacturer [12-13] both with regard to the stiffness and to the strength of the adhesives.

The maximum Rankine stress at failure in the SLJs (Fig. 9) highlights two main differences with respect to the one in the TBJs (Fig. 7). First, the failure stresses in the SLJs are many times higher than in the TBJs for both adhesives. Second, a remarkable decrease of the peak Rankine stress is evident not only for Multibond 330 but also for Hysol 9514. On the one hand, the higher failure stress that the SLJ can withstand can be attributed to its localization in a small amount of adhesive, due to their high gradient [23]. On the other hand, since a nearly-interface failure occurs with Hysol 9514, the failure of the SLJs

probably cannot be attributed to the structural stresses acting on the midplane of the adhesive.

Figure 10 and 11 provide a clear comparison between the peak stresses in the TBJs and the SLJs. Figure 10a (Multibond 330) and 11a (Hysol 9514) evidence a similar trend in the mean shear stresses at failure that originates in the joints, but a main difference appears. Multibond 330 (Fig. 10a) reaches higher shear stresses in the SLJ than in the TBJ, while Hysol 9514 (Fig. 11a) exhibits an opposite behaviour. This may be imputable to the higher ratio of the Young's modulus of the adherends to the one of the adhesive, occurring for Multibond 330. Figure 10b and 11b depict a slightly different behaviour between the adhesives with regard to the maximum shear stresses at failure. In the SLJ, Multibond 330 (Fig. 10b) exhibits a peak shear stress which is approximately twice the one in the TBJ, in the whole range of adhesive thickness. Hysol 9514 (Fig. 11b) shows the same ratio between the stresses as Multibond 330 only at the lower adhesive thickness, and then a noticeable disagreement between the trends is evident. By examining the maximum Rankine stresses at failure (Fig. 10c and 11c) similar observations can be drawn. In the SLJ, Multibond 330 (Fig. 10c) yields a Rankine stress about four times the one in the TBJ, in the whole range of adhesive thickness. The same ratio is obtained with Hysol 9514 (Fig. 11c), but only at the lower adhesive thickness. In the whole, the disagreement between the structural stresses in the TBJ and in the SLJ for the Hysol 9514 (Fig. 11b and c) supports the previous observation about the non applicability of such stresses to predict the failure of the SLJ.

By observing Tab. 6 it appears that the SIF is in substantial disagreement with the experimental static strength. The lower SIF at increasing thickness of the adhesive, in particular for Multibond 330, do not support the experimental lower strength registered in the SLJs. This marked contrast can be attributed to two main reasons. First, the noticeable yielding of the adherends observed in the experimental tests. Second, the small value of the adhesive layer thickness. As clearly described by Akisanya et al. [24], these conditions make the SIF approach not applicable to the examined configurations.

On the whole, when cohesive failure occurs, the proposed TBJ provides intrinsic static shear strength of the adhesive, coherent with the strength-thickness dependence experimentally observed in a SLJ. Regardless of the type of failure of the adhesive, the mean shear stress at failure (either torque over polar section modulus or force over area) appears to exhibit a good agreement between the TBJ and the SLJ.

5. CONCLUSIONS

This paper evaluates the effect of the adhesive thickness on the intrinsic static shear strength of the adhesive. Furthermore, the intrinsic strength, measured with an ad-hoc tubular butt joint (TBJ), is compared with the critical stresses at failure in a real bonded joint. The experimental tests compared the TBJ with a standard single lap joint (SLJ) by considering a range of the adhesive thicknesses from 0.05 mm to 0.4 mm. A modified methacrylate (Multibond 330), which exhibits a cohesive failure, and a high strength epoxy (Hysol 9514), which mainly fails close to the interface, are examined. Four observations can be drawn from the results presented in the work. First, a significant increase in the intrinsic static strength of the adhesive, as the adhesive thickness decreases, is observed when cohesive failure occurs (Multibond 330). By contrast, the strength is nearly independent of the adhesive thickness when interface failure occurs (Hysol 9514). Second, when cohesive failure occurs, the decreasing intrinsic shear strength as the adhesive

thickness increases is able to explain the decreasing strength experimentally observed in the SLJ for increasing adhesive thickness. Third, regardless of the type of failure, the average shear stress in the adhesive appears to exhibit a good agreement between the TBJ and the SLJ. Fourth, due to the yielding of the adherends and to the small thickness, the stress intensity factor cannot be applied to explain the experimental lower strength registered in the SLJs for increasing adhesive thickness.

REFERENCES

- [1] Dolev, G., and Ishai, O., 1981. "Mechanical Characterization of adhesive layer in-situ as bulk material", *J. Adhesion*, **12**, pp. 283-294.
- [2] Crocombe, A. D., 1989. "Global yielding as a failure criterion for bonded joints", *Int. J. Adhes. Adhes.*, **9**(3), pp. 145-153.
- [3] Gleich, D. M., Van Tooren, M. J. L., and Beukers, A., 2001. "Analysis and evaluation of bondline thickness effects on failure load in adhesively bonded structures", *J. Adhesion Sci. Technol.*, **15**(9), pp. 1091-1101.
- [4] Da Silva, L. F. M., Rodrigues, T. N. S. S., Figueiredo, M. A. V., de Moura, M. F. S. F., and Chousal, J. A. G., 2006. "Effect of adhesive type and thickness on the lap-shear strength", *J. Adhesion*, **82**(11), pp. 1091-1115.
- [5] Goglio, L., Rossetto, M., and Dragoni, E., 2008. "Design of adhesive joints based on peak elastic stresses", *Int. J. Adhes. Adhes.*, **28**, pp. 427-435.
- [6] Goglio, L., and Rossetto, M., 2008. "Impact rupture of structural adhesives joints under different stress combinations", *Int. J. of Impact Eng.*, **35**, pp. 635-643.
- [7] Goglio, L., and Rossetto, M., 2010. "Stress intensity factor in bonded joints: Influence of the geometry", *Int. J. Adhes. Adhes.*, **30**, pp. 313-321.
- [8] Adams, R. D., and Peppiatt, N. A., 1974. "Stress analysis of adhesive-bonded lap joints", *J. Strain Analysis for Engng. Des.*, **9**(3), pp. 185-196.
- [9] Adnan, A., and Sun, C. T., 2008. "Effect of adhesive thickness on joint strength: a molecular dynamics perspective", *J. of Adhesion*, **84**(5), pp. 401-420.
- [10] Dragoni, E., and Mauri, P., 2000. "Intrinsic static strength of friction interfaces augmented with anaerobic adhesives", *Int. J. of Adhesion and Adhesives*, **20**(4), pp. 315-321.
- [11] In book, Adams, R. D., Comyn, J., and Wake, W.C., 1997. *Structural adhesive joints in engineering*, 2nd ed. Springer, London, Chap. 3, pp. 127-128.
- [12] Tech Datasheet, 2007. Loctite – Multibond 330, November.
- [13] Tech Datasheet, 2003. Loctite – Hysol 9514, July.
- [14] ABAQUS 6.7, 2007. Users' Manual, HKS Inc.
- [15] EN 10083-3: Steels for quenching and tempering. Technical delivery conditions for alloy steels, 2006.
- [16] ASTM A-109: Standard Specification for Steel, Strip, Carbon (0.25 Maximum Percent), Cold-Rolled, 2008.

- [17] Peroni, L., Avalue, M., and Belingardi, G., 2009. "Comparison of the energy absorption capability of crash boxes by spot-weld and continuous joining techniques", *Int. J. of Impact Eng.*, **36**, pp. 498-511.
- [18] Castagnetti, D., and Dragoni, E., 2009. "Standard finite element techniques for efficient stress analysis of adhesive joints", *Int. J. of Adhes. and Adhes.*, **29**, pp. 125-135.
- [19] Castagnetti, D., Dragoni, E., and Spaggiari, A., 2009. "Efficient post-elastic analysis of bonded joints by standard finite element techniques", *J. of Adhes. Sci. and Tech.*, **23**, 1459-1476.
- [20] Castagnetti, D., Dragoni, E., and Spaggiari, A., 2010. "Failure analysis of bonded T-peel joints: efficient modelling by standard finite elements with experimental validation", *Int. J. of Adhes. And Adhes.*, **30**, pp. 306-312.
- [21] Wang, C. H., and Rose, L. R. F., 2000. "Compact solutions for the corner singularity in bonded lap joints", *Int. J. of Adhesion and Adhesives*, **20**, pp. 145-154.
- [22] Pirondi, A., and Moroni, F., 2009. "An investigation of Fatigue Failure prediction of adhesively bonded metal/metal joints", *Int. J. of Adhes. and Adhes.*, **29**, pp. 796-805.
- [23] Bigwood, D. A., and Crocombe, A. D., 1989. "Elastic analysis and engineering design formulae for bonded joints", *Int. J. of Adhesion and Adhesives*, **9**, pp. 229-242.
- [24] Akisanya, A. R., and Meng, C. S., 2003. "Initiation of fracture at the interface corner of biomaterial joints", *J. of the Mech. of Physics and Solids*, **51**, pp. 27-46.

USE OF NEURAL NETWORKS IN CEREBRAL PALSID CHILDREN: FORM RECOGNITION BY MEANS OF GAIT ANALYSIS

Giacomo Reggiani

Department of Science and Engineering Methods,

University of Modena and Reggio Emilia, Italy

E-mail: 60345@studenti.unimore.it

Marco Cocconcelli

InterMech MoRE, Department of Science and Engineering Methods,

University of Modena and Reggio Emilia, Italy

E-mail: marco.cocconcelli@studenti.unimore.it

Riccardo Rubini

Department of Science and Engineering Methods,

University of Modena and Reggio Emilia, Italy

E-mail: riccardo.rubini@studenti.unimore.it

Corrado Borghi

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: corrado.borghi@gmail.com

Alberto Ferrari

LAMBDA gait analysis laboratory,

Azienda Ospedaliera Arcispedale S. Maria Nuova, Reggio Emilia, Italy

E-mail: alberto.ferrari8@gmail.com

Abstract: *The classification systems for cerebral palsy (CP) need to be continuously updated, according to specific objectives and to significant changes observed over the years in the panorama of CP. Ferrari et al. [1], recently proposed a classification system that aimed at subdividing the diplegic children into four main clinical sub-forms, on the base of their walking pattern. This preliminary study deals with the classification of the diplegic children affected by CP and it is based on the walking pattern classification system proposed by Ferrari and utilized in LAMBDA motion analysis laboratory at S. Maria Nuova Hospital of Reggio Emilia. Using kinematics data recorded by means of an optoelectronic system on children affected by CP, an Artificial Neural Network (ANN) was implemented to allow an automatic recognition of the form of the palsy. The ANN proposed correlates a set of suitable statistical parameters of the kinematics of walking with the type of diplegic clinical form. The effectiveness of the resulting neural network has been proved on a control set of data.*

Keywords: *Artificial Neural Networks, Cerebral palsy, Gait analysis*

1. INTRODUCTION

This study aims at supporting the work of clinical specialists (i.e. physicians and physical therapists) caring for children affected by Cerebral Palsy (CP). This disorder is characterized by chronic impairment of posture and movement due to an early occurrence of a stable lesion to the brain [1]. Several classifications of the forms of CP have been historically proposed. The importance of a classification system stays in the description that it can give to the disorder. A good classification allows the operators to hypothesize the “natural history” a young patient is going to follow along his/her growth on the base of other patients with similar clinical signs, therefore to declare a priori the site and the measure of the modifiability of his/her adaptive functions [1] and the best result reachable from therapy. In other words to define the functional limits imposed by the lesion and to predict the evolution of the motor abilities on the base of the disorder. A good classification system is unequivocally necessary to orient and optimize the therapy. This work is based on the walking pattern classification system proposed by Ferrari et al. [1] and used in LAMBDA motion analysis laboratory at S. Maria Nuova Hospital of Reggio Emilia, and aimed at defining a method to perform automatically and objectively this classification system.

In particular, the classification system proposed by Ferrari, separates the macro category “diplegia” into four main subgroups, or forms. The assignment of a patient to a form is currently obtained by visual observation from expert operators. In order to demonstrate the reliability of the classification and to spread its use all over the world, in S. Maria Nuova Hospital specific technical tools were adopted and a gait analysis (GA) laboratory was set up (see paragraph “Gait analysis data” for more details).

The goal of this work is to define an automatic method based on data collected from GA to objectively classify a patient according to the Ferrari classification.

One of the most reliable methods applied in computer science to automatically classify cluster of data is artificial intelligence and in particular Supervised Learning Approaches. In this field Artificial Neural Networks (ANN) have been proficiently used to solve many problems especially thanks to their capability in managing a large number of input variables [4-6].

In particular, ANN have been successfully applied in mechanics, in order to diagnose machines usury status by means of vibrational signals [7], and in the automotive sector, for road recognition systems [8-9].

ANN have been also already applied in medicine for diagnostic purposes, for cancer in particular. In medicine the diagnostic activity, indeed, often cannot be reduced to a linear problem and requires a wide spectrum of input variables. For example ANN are suitable to classify a tumor both benign or malignant, on the base of cell scansion gathered by microscopic examination, or to predict cholesterol levels (ldl, hdl and vldl), on the base of 21 different spectral components. Furthermore, it is also possible to use ANN to detect diabetes, exploiting daily life physiologic data [10].

For what concern human kinematics, Bayesian networks have been recently used to classify the walking pattern of C.P. children and have predictive value on the degree of crouch and stiff knee [11].

In this paper a new application of ANN is proposed. The method is based on the use of statistical parameters collected from the gait kinematics of the patient as recorded by means of a GA exam with the aim of identifying the affiliation to one of the four diplegic forms as described in [1]. From the raw kinematics data of the walking pre-processing algorithms were applied in order to obtain variables meaningful for the classification, that is to

translate in objective parameters the clinical signs used by medical doctors to perform the classification. These variables and parameters were then used in order to train different structures of ANN.

In this paper it is first described how GA can be used for pattern classification purposes, then the protocol used to reproduce automatically the classification of Ferrari. Subsequently, the pre-processing of signals and the organization of the parameters required by the ANN are reported. Finally the main results obtained on the base of the cluster of patients acquired so far in the laboratory of Reggio Emilia are showed.

2. THE FOUR MAIN CLINICAL FORMS OF CP

The treatment of the CP is considered as the “core problem” for the whole field of the pediatric rehabilitation. The reason why this pathology has such a primary role, can be ascribed to two main aspects. First of all CP is the form of disability most frequent in childhood (one new case per 500 birth alive, [1]), secondarily the functional recovery of the “spastic” child is, historically, the clinical field in which the majority of the therapeutic methods and techniques (physiotherapy, orthotic, pharmacologic, orthopedic-surgical, neurosurgical) were first applied and tested. The currently accepted definition of CP – Group of disorders of the development of movement and posture causing activity limitation [2] – is the result of a recent update by the World Health Organization to the language of the International Classification of Functioning Disability and Health, from the original proposal of Ingram – A persistent but not unchangeable disorder of posture and movement – dated 1955 [3]. This definition considers CP as a permanent ailment, i.e. a “fixed” condition, that however can be modified both functionally and structurally by means of child spontaneous evolution and treatments carried out during childhood. The lesion that causes the palsy, happens in a structurally immature brain in the pre-, peri- or post-birth period (but only during the first months of life).

The lesion, whether focused or spread over the nervous system, impairs the whole functioning of the Central Nervous System. As a consequence, it affects the construction of the adaptive functions [4], first of all posture control, locomotion and manipulation.

However, in the scientific and clinic community no common classification system of CP has so far been universally accepted. Besides, no standard operative method or technique have been acknowledged to effectively assess the different disabilities and impairments exhibited by children with CP. CP is still “an artificial concept, comprising several causes and clinical syndromes that have been grouped together for a convenience of management” [5].

Studying for more than twenty years the development of adaptive functions from a huge sample of CP children, coming from all over Italy and distributed over a wide range of age (from a few months to adulthood), the group leaded by Ferrari Adriano operating in S. Maria Nuova Hospital in Reggio Emilia has been able to identify and classify the recurrent pathognomic motor patterns characterizing the status of the motor disabilities expressed by children with diplegic CP. These pathognomic signs are therefore currently favorably used to classify CP in different forms in the centre of Reggio Emilia and many others across Italy.

The CP classification system proposed by Ferrari et al. [1], divides the spastic forms of diplegia according to the architecture of a guiding function: the walking. Indeed, the architecture of walking contemporarily provides the guiding criterion for the diagnosis of clinical form and many useful indications for prognosis and therapeutic decisions. Moreover walking represents for diplegic CP children, since they have different patterns

and are not always achievable spontaneously or maintainable during the adulthood, the activity to which the main therapeutic engagement (physiotherapy, orthosis, drugs, functional surgery) is directed.

The classification separates the macro category “diplegia” into four main subgroups, in which the architecture of walking function shows meaningful clinical and prognostic differences. In practice the following main elements, or signs, are considered:

- use of the upper limbs and walking aids;
- attitude of head and trunk;
- pendular movements of the trunk in sagittal and frontal planes;
- movements of the pelvis;
- mechanism of progression;
- sequence of strike and foot off.

Table (1) summarizes the different patterns and pathognomic signs exhibited by the four different forms. The subdivision from a generic group into homogeneous clusters drives the operators during the clinical decision making process in designing an accurate and patient-oriented rehabilitation plan.

In order to spread the use of this classification system in other clinical centers and to increase its robustness, instruments able to measure the gait pattern could be used. In particular, the most promising tool that can be applied to perform the classification is definitely the technique known as gait analysis.

3. GAIT ANALYSIS

Analysis of human posture and movement is a fast-growing biomedical sector and matter of great interest from the clinical point of view. Since the ‘80s, instruments specifically oriented to the analysis of the human movement have been advantageously designed and applied in the context of CP with the aim of measuring motor deficits and, especially, gait deviations. The GA technique has been increasingly used over the years to assess, analyze, classify, and support the process of clinical decisions making, allowing for a complete investigation of gait with an increased temporal and spatial resolution.

Table 1. Forms and their main signs.

	I form “forward leaning propulsion”	II form “tight skirt”	III form “tightrope walker”	IV form “dare devils”
Trunk	bending forward	vertical	bending forward	vertical
Pendulum	sagittal	sagittal - frontal	frontal	counter shoulders/pelvis
Progression	rotation with respect to the hip	flexion with respect to the knee	foot pivot	hip intrarotation
Feet	contact equinus	contact equinus / clubfoot	starting/ pushing equinus	starting/ pushing equinus

GA has provided a basis for improving the outcome of surgical and nonsurgical treatments and for introducing a new *modus operandi* in the identification of defects and functional adaptations to the musculoskeletal disorders. A widespread application of reliable instruments and techniques able to objectively evaluate both the form of the palsy (diagnosis) and the efficacy of the treatments provided (prognosis), constitutes a valuable method able to validate care protocols, establish the efficacy of classification systems and assess the validity of definitions.

The laboratory of S. Maria Nuova Hospital (LAMBDA, Laboratorio di Analisi del Movimento del Bambino Disabile) utilized for this study is equipped with the following instruments: optoelectronic system, environmental cameras, force platforms, electromiographic (EMG) system.

The optoelectronic system detects the three-dimensional (3D) coordinates of reflective markers fixed to the body of the subject. The markers are illuminated by infra-red light. Each of the eight cameras of the laboratory identify the markers reflexes on a 2D camera reference frame. These data, conveniently processed, permit to identify the position of the markers on the 3D laboratory reference frame. This system is therefore able to reconstruct the walking kinematics of the subject, in particular the angles of flexion-extension, abduction and in-external rotation at the main joints of the lower limbs: hip, knee and ankle.

Two environmental cameras generate movies for the visual observation of the subject after the acquisition. Two force platforms, positioned at the floor level, measure torques and forces that the subject exerts to the ground. Coupling this measure and the kinematics data makes it possible to calculate moments and powers at the joints level.

The EMG system finally assess the muscles activation pattern during the gait.

A view of the LAMBDA laboratory is shown in Fig. (1).

4. DESCRIPTION OF THE PROBLEM

The raw data collected from an optical system, as the one installed in LAMBDA, consist in trajectories of markers positioned in particular points of the body. By means of the use of a biomechanical model specifically designed to obtain relevant kinematic information, is possible from the trajectories of the markers to represent the walking kinematics as joint angles and joint moments, and effectively report defects and compensations of a pathological scheme by showing these variables as function of a gait cycle. These data refers both to the upper and lower part of the body. The biomechanical model simply foreseen to position the markers in a priori established relevant anatomical points, known as landmarks.

In order to obtain the clinical relevant parameters used by operators during the classification of the diplegic forms process, pre-processing computations were performed especially based on the dynamic of particular distances between different parts of segments of the body. In fact this relative displacements represent the variation of position of anatomical joints during walking.

The evaluated forms and their representative characteristics are reported in Tab. (1).

Due to the fact that the technology and the protocols used in the context of rehabilitation of children with CP are usually standard and aligned with the one used in LAMBDA, the ANN developed is feasible to be installed as separate plug-in into many others laboratories of gait analysis.

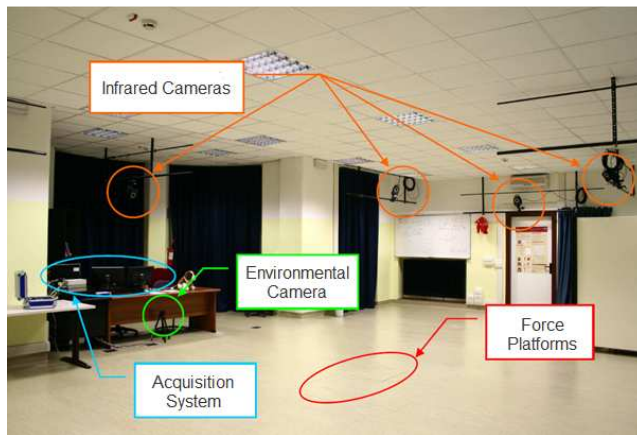


Figure 1. LAMBDA laboratory with its equipment.

5. PRE-PROCESSING OF THE KINEMATICS

Experimental tests supplied 57 patients not equally divided into the four forms of CP. Each patient was request to repeat several trials in order to increase the number of recording. Because of the differences in the height of the patients and the speed of the walking performed, the number of gait cycles for each trial is different varying from 2 to 4. The total amount of trials, consisting in walks of 2-4 gait cycles, is 283. Patients of the third and fourth form, that show the highest motor abilities, are able to execute more steps in the same trial compared to those of first and second forms. In order to standardize the data sets, only three cycles of steps were considered for each trial recording.

Because of different possible direction of walk, in order to refer parameters just with respect to the frontal and the sagittal plane of the patient, the least square line of the marker positioned on L5, which easily represent the walking direction, was calculated for each trial acquisition.

The total amount of records, obtained by the described cutting procedure, was separately processed to extrapolate mechanical and, secondarily, statistical parameters. The first ones evaluate instantaneous information about patient kinematics that are then used in order to derive the statistical values linked to the global properties of the form.

Mechanical data consists of many distances between particular anatomical points of the upper and lower part of the body calculated along the three cycles of steps and normalized with respect to the length of the segment joint registered during a static acquisition with the subject standing in upright posture. In particular, the upper and lower limbs were normalized with respect to the legs length. Each anatomical position considered was analyzed in the laboratory reference coordinate system that sets a XY plane for the floor and Z axis for the vertical direction.

The upper limbs landmark considered were *acromion* (xA), *radius epicondyle* (xEP) and *ulna* (xUL) of each arm (x represents the right or left side). Then the distances from acromion to ulna ($x_A - x_{UL}$) were calculated separating the components in the XY plane and in the Z direction.

The trunk was represented by the distance from the *7th cervical vertebra* (C7) to the *5th lumbar vertebra* (L5) calculated along the direction of walking and perpendicularly to that line.

To detect the kinematic of the pelvis we estimated the distance between *axis* (xASIS) and the *lateral epicondyle of the knee*, measured in the vertical direction.

The lower limbs were schematize by the length from the *great trochanter* (xGT) to the *lateral malleolus of the ankle* (xLM) for the entire leg, considered along the Z axis, while the position of feet referred to the leg is represented by the distance between the *tibial tuberosity* (xTT) and the *1st metatarsal* (xFM) on the XY plane.

The exact position of markers considered is reported in Fig. (2), while the kinematics signal are summarized in Tab. (2).

Because of huge dimension of the input vector all the distances calculated were summarized with four statistical parameters in order to facilitate the task for the network.

The firsts statistical operators selected as input parameters for the network are the mean value and the variance of the distance kinematic signals, that summarize the kinematics information in one number representative of the order of magnitude and describe the main tendency of the displacements of the markers respect to the mean value. In addition we also use the skewness and kurtosis that measure the deviation from the normal distribution.

Therefore each acquisition is represented by an input vector made up of 48 elements: 4 statistical parameters of the twelve distances considered.

For each input array, a binary output is joined to relate the CP's form to the correspondent patient kinematics: the 4 typologies are identified by an array constituted by 4 binary elements.

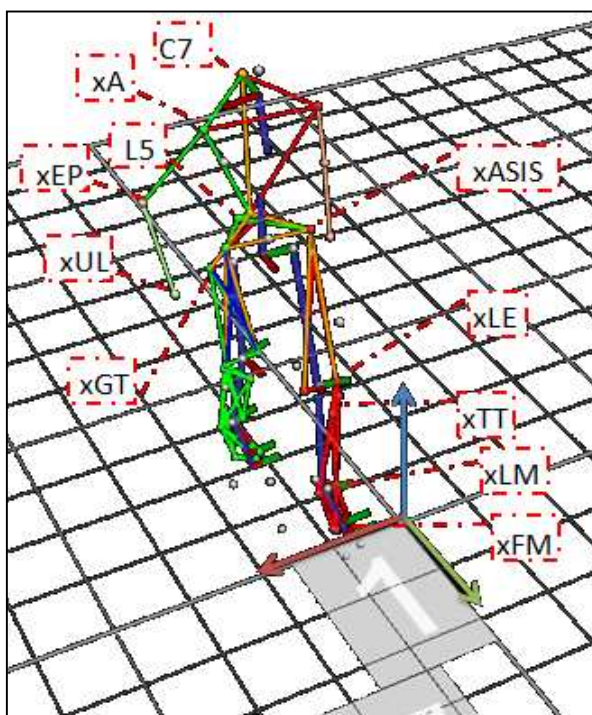


Figure 2. Schematic representation of markers position and laboratory reference system (Red, Green and Blue vectors corresponds to X, Y and Z axis)

Table 2. Main distances (“x” means the left or right side of the body).

SEGMENTS	DISTANCES
UPPER LIMBS (xA – xUL)	Z direction
	XY plane
TRUNK (C7 – L5)	Direction of walking
	Perpendicular to the direction of walking
PELVIS (xASIS - xLE)	Z direction
BOTTOM LIMBS (xGT – xLM)	Z direction
FEET (xTT – xFM)	XY plane

Part of the total set of records was assigned to validate the neural network, in order to highlight its performance. For this reason, a test campaign was managed with different amount of training records respect to the global acquisitions: 100/283, 154/283 and 199/283.

The data used for validation was not in common with the ones used for training, and the results of the consequent network simulations were then compared to the target output from the clinical assessment. So we reserved a part of the data of each form for the performance estimation only. Since the output array is included between 0 and 1 while the target is a binary array, the elements of the output array are normalized with respect to the maximum value and then rounded to the floor.

6. RESULTS

The neural network used is a *Feed-Forward* architecture which is the most used in the *Pattern Recognition* problems [11-13]. The neural network is implemented with the dedicated Matlab toolbox: since the input array size is greater than the output one, the possible architectures are *convergent* and *divergent-convergent*; as a matter of fact, the network configuration depends on the number of neurons of the intermediate layers. A number of hidden layers included between 1 and 3 was tested, with several different numbers of neurons for each layer.

All the architectures were trained with the three different subdivision of data set: 100 acquisition reserved for the training procedure, 154 and 199.

Each configuration was going to be analyzed varying some parameters of the Neural Network such as the transferring function or the training function.

Evaluated transferring functions were *tansig* and *logsig*, while only the training functions suitable for the problems of pattern recognition were considered: the algorithm of resilient back-propagation (*trainrp*), the algorithm of Levenberg-Marquardt (*trainlm*) and the algorithm of scaled conjugate gradient (*trainscg*).

All the mentioned transferring functions and training functions were implemented in Matlab Neural Network Toolbox [10].

The difference between transferring functions *tansig* and *logsig* is showed in Fig. (3). *Tansig* output lies in the interval [-1; +1], while the *logsig* output in [0; +1]. For further information about training functions refer to [13].

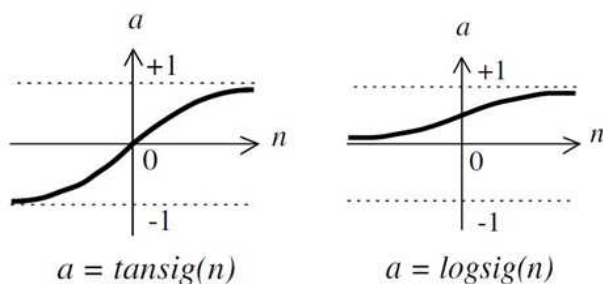


Figure 3. Transferring functions in Matlab Neural Network Toolbox

In order to find the more effective setting parameters and to estimate the robustness of the network, only one parameter at a time was modified. Ten simulations per type of network were executed and statistically managed regarding in particular the error percentage.

Table (3) reports the error percentage for the best type of neural network considered versus the number of acquisitions using during the training fase. The configuration chosen as reference one has the following characteristics:

- Training function: *trainrp*
- Transfer function: *tansig*

The best performance is highlighted in yellow.

Analyzing the results obtained after the above mentioned simulations, it is possible to suggest which architecture can further improve the network performances:

- divergent-convergent architecture
- high number of hidden layers
- high number of neurons
- transferring function *tansig*
- training function *trainscg*

The optimum tested architecture is constituted by 3 hidden layers, with respectively 4, 8 and 6 neurons, transferring function *tansig*, learning function *trainscg* and a an amount of learning data equal to 199 records of the total acquired, whose schematic representation is proposed in Fig. (4). These settings provide a neural network with a percentage of correct input recognition equal to 63%, i.e. 53 samples were perfectly recognized over 83 proposed to the network.

Table 3. Percentage of errors changing the amount of input vector in the training.

ARCHITECTURE	ERROR PERCENTAGE		
	Data Subdivision		
	100/283	154/283	199/283
3 Layers 4/8/6	57%	49%	37%

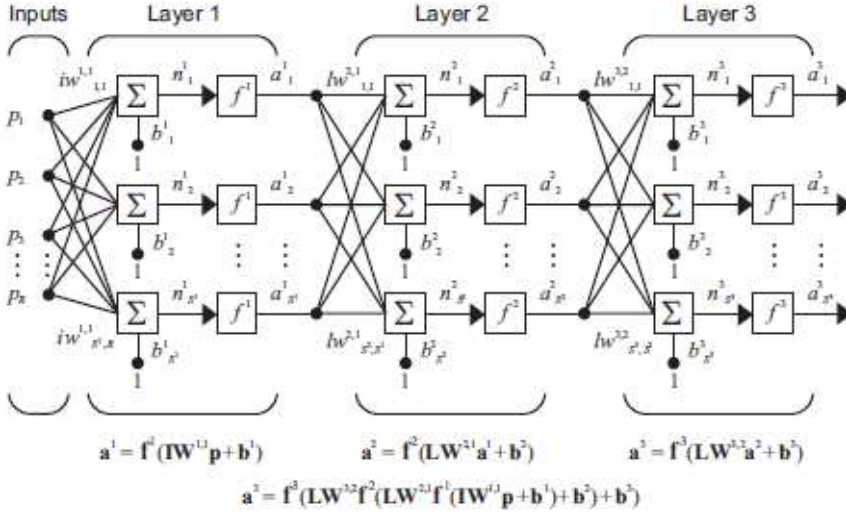


Figure 4. Schematic representation of the best network architecture

7. CONCLUSIONS

In this study kinematic signals representative of the gait pattern collected from 57 CP patients belonging to one of the four clinical forms described by Ferrari et al. [1], were registered by an optoelectronic system in the tridimensional space of the LAMBDA laboratory. An automatic ANN algorithm able to classify patients in the four different forms was developed and tested. The mean value, variance, skewness, and kurtosis from relevant kinematic variables were calculated on the base of the raw data represented by the distances between specific anatomical landmarks. The resulting cluster of data for the ANN was of 283 input vectors constituted by 48 elements. Test vectors were composed by 4 element of zero values and one unitary element determining the diplegic form. Many convergent and divergent-convergent architectures, with different number of neurons and hidden layers, were tested dividing the data set in different amounts of acquisitions (100/183, 154/129, 199/84): the first part was used for training the neural network, the second for testing the results. This type of subdivision allows to estimate the performance of the network regardless of data used for training, giving more accurate results.

The most reliable and efficient architecture was tested also changing the neuronal transferring function and the training function of the general network.

Despite the success rate is far from acceptable values of the methodology in the clinical results are encouraging.

It is well known from literature that the performance increases with the number of training samples for the network. To have reliable results, the average number of samples should be four times higher than the current database available in LAMBDA laboratory.

Future developments and modifications of the kinematical and statistical parameters used as input to the network are needed in order to improve performance and usability of the algorithm. This work is object of a future journal article.

Acknowledgements

This paper is supported by the Tecnopolo InterMech-MoRE, Italy.

REFERENCES

- [1] Leardini, A, Sawacha, Z, Paolini, G, Ingrosso, S, Nativo, R, and Benedetti, MG., 2007, "A new anatomically based protocol for gait analysis in children", *Gait Posture*, **26**(4), Oct, pp. 560-571.
- [2] Ferrari, A., and Cioni, G., 2005, *Le forme spastiche della paralisi cerebrale infantile*, Springer
- [3] Ingram, TT., 1955, "A study of cerebral palsy in the childhood population of Edinburgh", *Archives of Disease in Childhood*, **30**(150), Apr, pp. 85-89.
- [4] Nilsson, N. J., 2005, Introduction to Machine Learning. See also URL <http://ai.stanford.edu/~nilsson/mlbook.html>
- [5] Fausett, L., 1994, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*, Prentice Hall
- [6] Kriesel, D., 2007, A Brief Introduction to Neural Networks. See also URL: http://www.dkriesel.com/en/science/neural_networks
- [7] Giuliani, G., Rubini, R., and Maggiore, A., 1998, "Ball Bearing Diagnostics Using Neural Networks", *Proceedings of the Third International Conference Acoustical and Vibratory Surveillance Methods and Diagnostic Techniques*, Senlis, France, pp.767-776.
- [8] Reggiani, G., Cocconcelli, M., Rubini, R., and Lolli, F., 2010, "Use of Neural Networks in Road Recognition by Vibration Data", *Proceeding of the ASME 2010 International Design Engineering Technical Conference & Computer and Information in Engineering Conference IDETC2010*, August 15-18, 2010, Montreal, Quebec, Canada.
- [9] Tanovic, O., Huseinbegovic, S., and Lacevic, B., 2008, "Road Type Recognition Using Neural Networks for Vehicle Seat Vibration Damping", *Proceedings of the IEEE International Symposium on Signal Processing and Information Technology*, December 16-19, Sarajevo, Bosnia and Herzegovina pp.320-323.
- [10] Demuth, H., Beale, M., and Hagan, M., 2008, *Neural Network Toolbox 6: User's Guide*.
- [11] Schwarts, M., et al., 2010, "Classification of Gait Patterns with an emphasis on Cerebral Palsy" *Proceedings of the Clinical Gait Analysis Course: a focus on interpretation. 20th annual conference*, July 1-3, 2010 Leuven, Belgium.
- [12] Nigrin, A., 1993, *Neural networks for pattern recognition*, Bradford Books
- [13] Bishop, C. M., 2004, *Neural networks for pattern recognition*, Oxford University Press.
- [14] Sandberg, I.W., Lo, J.T., Fancourt, C.L., Principe, J.C., Katagiri, S., and Haykin S., 2001, *Nonlinear Dynamical Systems: Feedforward Neural Network Perspectives*, John Wiley & Sons.

ACCELEROMETRI MEMS: CARATTERIZZAZIONE DINAMICA E CONFRONTO CON I SENSORI PIEZOELETTRICI

Giovanni Scirè Mammano

Dipartimento di Scienze e Metodi dell'Ingegneria,

Università di Modena e Reggio Emilia, Italia

E-mail: giovanni.sciremammano@unimore.it

Marco Cocconcelli

InterMech MoRE, Dipartimento di Scienze e Metodi dell'Ingegneria,

Università di Modena e Reggio Emilia, Italia

E-mail: marco.cocconcelli@unimore.it

Riccardo Rubini

Dipartimento di Scienze e Metodi dell'Ingegneria,

Università di Modena e Reggio Emilia, Italia

E-mail: riccardo.rubini@unimore.it

Eugenio Dragoni

Dipartimento di Scienze e Metodi dell'Ingegneria,

Università di Modena e Reggio Emilia, Italia

E-mail: eugenio.dragoni@unimore.it

Abstract. *Gli accelerometri di tipo piezoelettrico sono probabilmente i più diffusi sensori di accelerazione attualmente in commercio per l'ambito industriale. Hanno infatti un'ampia banda dinamica di linearità, sono influenzati in modo trascurabile da rumore e temperatura, e possono misurare elevati valori di accelerazione. Negli ultimi anni sono stati proposti sul mercato dei nuovi tipi di accelerometri di tipo MEMS (Micro Electro-Mechanical Systems) che hanno la caratteristica di essere molto meno costosi dei piezoelettrici sia per i materiali impiegati sia per il processo di fabbricazione di massa. Se fino a poco tempo fa la qualità degli accelerometri MEMS non era paragonabile a quella dei piezoelettrici, la continua evoluzione tecnologica ha estremamente ridotto le differenze tra le due tipologie di sensori. Lo scopo di questo articolo è di confrontare le prestazioni dinamiche di alcuni accelerometri di tipo piezoelettrico e di tipo MEMS, per valutarne le differenze, qualità e difetti, e fornire un'introduzione alle problematiche che nascono nel passaggio da una tecnologia all'altra.*

Keywords: *Accelerometri MEMS, Accelerometri piezoelettrici, Dinamica*

1. INTRODUZIONE

In ambiente industriale il monitoraggio delle macchine è utile al fine di salvaguardare l'integrità strutturale delle macchine stesse o di un insieme di componenti quali ad esempio ingranaggi e cuscinetti. L'individuazione di difetti meccanici, che sono la principale causa

di rottura, è possibile grazie all'acquisizione di segnali vibratorii da parte di accelerometri. Attualmente gli accelerometri piezoelettrici sono probabilmente i sensori più diffusi per la misura delle vibrazioni in quanto possono essere utilizzati in campi di frequenze molto ampi mantenendo una buona linearità e accuratezza anche in condizioni ambientali gravose a fronte di un costo del sensore che ne limita l'impiego.

Negli ultimi anni è stata proposta sul mercato una nuova tipologia di accelerometri sviluppati grazie alla tecnologia MEMS (Micro Electro Mechanical System). Questi innovativi accelerometri sono di dimensioni e peso ridotti e poiché possono essere fabbricati in grandi quantità risultano essere a basso costo. Richiedono un limitato consumo di potenza e consentono un elevato livello di integrazione tra parte meccanica ed elettronica ovvero in un unico processore convivono sia il sensore di misura sia il circuito elettronico di condizionamento del segnale. Il recente successo degli accelerometri MEMS è dovuto al loro impiego in prodotti di largo consumo come cellulari, airbag, GPS e controller di videogiochi, in realtà è sin dagli anni novanta sono in atto studi approfonditi per la loro applicazione. Ad esempio la NASA ha pensato per prima alla possibilità di inserirli in programmi spaziali per ridurre i costi delle missioni. Alcuni studi dimostrano che la robustezza meccanica di questi accelerometri è adeguata per la maggior parte delle applicazioni aerospaziali [1].

Il presente studio si propone di valutare le reali capacità degli accelerometri MEMS limitatamente al campo della diagnostica delle macchine. Lo scopo dello studio è capire se gli accelerometri MEMS sono in grado di effettuare misure di vibrazioni affidabili come quelle degli accelerometri piezoelettrici per una banda di frequenze di interesse per il monitoraggio dei componenti meccanici. Obiettivo della ricerca è ottenere misure di accelerazione corrette utilizzando sensori a basso costo.

L'articolo è strutturato in quattro paragrafi: nel primo vengono presentate le tipologie di sensori, le loro principali caratteristiche e i principi di funzionamento; in quello successivo l'attrezzatura sperimentale, le caratteristiche da catalogo degli accelerometri utilizzati e le prove eseguite per caratterizzarli dinamicamente. Nel terzo paragrafo gli accelerometri vengono utilizzati per effettuare l'analisi modale di una lamina piana e testare quindi sperimentalmente il loro impiego. L'ultima parte include i commenti e le conclusioni dell'attività svolta.

2. TIPOLOGIE DI ACCELEROMETRI

Gli accelerometri possono essere classificati in base alla loro differente tipologia costruttiva che può essere [2]:

- Piezoresistiva
- Capacitiva
- Piezoelettrica

In questo articolo tratteremo le due tipologie più utilizzate che sono quella capacitiva per la produzione di accelerometri MEMS e la tipologia piezoelettrica., delle quali verranno descritti i principi di funzionamento.

Accelerometri Capacitivi

La maggior parte dei dispositivi MEMS sono fabbricati utilizzando la tipologia costruttiva capacitiva che permette una elevata sensibilità del sensore e una sostanziale indipendenza del valore della temperatura.

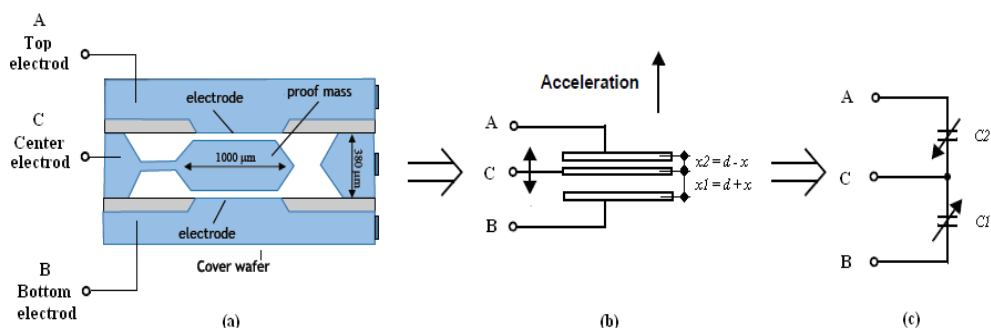


Figura 1. Accelerometro MEMS di tipo capacitivo e relativo schema elettrico.

I sensori capacitivi offrono eccellenti prestazioni per quanto riguarda il rapporto segnale/rumore e richiedono bassa potenza. Allo stato di sviluppo tecnologico attuale gli accelerometri MEMS capacitivi operano in un campo di accelerazione di centinaia di g e con frequenze fino a 1kHz [3].

La tecnologia capacitiva si basa sulla misura della variazione di capacità a fronte di un movimento meccanico. La Fig. (1a) riporta un accelerometro MEMS, in cui una massa a sbalzo viene a costituire il polo centrale oscillante comune a due condensatori. L'accelerazione è dedotta misurando la variazione di capacità tra la massa mobile e il package.

In particolare tra il “wafer” di copertura e la massa è possibile individuare due capacità (C_1 e C_2) come mostrato dallo schema elettrico in Fig (1b, 1c) esse variano il loro valore a seconda del movimento della massa.

Il valore della capacità di un condensatore a lamine parallele si può esprimere come segue:

$$C = \epsilon \frac{A}{d} = \epsilon_0 \epsilon_r \frac{A}{d} \quad (1)$$

dove A indica l'area delle lamine, d la distanza tra esse, ϵ_r la costante dielettrica relativa del mezzo interposto tra le lamine e ϵ_0 la costante dielettrica del vuoto.

A causa del movimento della massa sottoposta ad accelerazione il valore istantaneo delle capacità è:

$$C_1 = \epsilon \frac{A}{x_1} = \epsilon \frac{A}{d+x} = C_0 - \Delta C \quad (2)$$

$$C_2 = \epsilon \frac{A}{x_2} = \epsilon \frac{A}{d-x} = C_0 + \Delta C \quad (3)$$

dove $C_0 = \varepsilon \frac{A}{d}$ è il valore delle capacità quando la massa è ferma;

d è la distanza tra i piatti delle capacità quando la massa è ferma;

x è lo spostamento della massa, ovvero lo spostamento dei piatti dalla posizione assunta quando la massa è ferma.

Se l'accelerazione è nulla, la massa non oscilla e x vale zero. In questo caso le capacità C_1 e C_2 sono uguali. In caso contrario lo spostamento della massa è pari ad x e la differenza tra le capacità è:

$$C_2 - C_1 = 2\Delta C = 2\varepsilon A \frac{x}{d^2 - x^2} \quad (4)$$

Questa può essere riscritta come:

$$\Delta C \cdot x^2 + \varepsilon A \cdot x - \Delta C d^2 = 0 \quad (5)$$

Per piccoli spostamenti di x il termine di secondo grado $\Delta C \cdot x^2$ è trascurabile. La soluzione diviene:

$$x \cong \frac{\Delta C d^2}{\varepsilon A} = d \frac{\Delta C}{C_0} \quad (6)$$

La formula mette in luce che lo spostamento x è proporzionale alla variazione di capacità ΔC .

Il segnale deve essere prelevato da un circuito elettronico: una tipica configurazione è quella rappresentata in Fig (2) (ponte di Wheatstone capacitivo).

Poiché sul ramo sinistro del ponte vi sono due capacità di riferimento identiche di valore pari a C_0 la tensione V_1 vale:

$$V_1 = \frac{V_{in}}{2} \quad (7)$$

La tensione V_2 vale invece:

$$V_2 = \frac{C_2}{C_1 + C_2} V_{in} \quad (8)$$

La tensione misurata in uscita è:

$$V_{out} = V_2 - V_1 = \frac{x}{2d} V_{in} \quad (9)$$

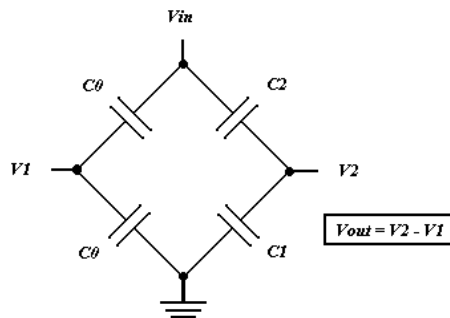


Figura 2. Ponte di Wheatstone capacitivo.

Grazie a questa configurazione la tensione in uscita è una funzione lineare dello spostamento x .

Il principio su cui si basa la tecnologia capacitiva è semplice ma spesso non è banale effettuare una misura dell'accelerazione in quanto la variazione di capacità è estremamente piccola. Ad esempio per accelerometri mostrati in Fig. (1) la variazione di capacità si attesta sui *femtofarad* per accelerazioni di un *milli-g*. [4].

Accelerometri Piezoelettrici

Questi sensori basano il loro funzionamento sul principio fisico noto come piezoelettricità secondo il quale alcuni materiali generano un campo elettrico se sottoposti ad una deformazione meccanica.

La tipologia costruttiva piezoelettrica mostra il vantaggio di essere *self-generating* in quanto, a differenza di quella capacitiva, non necessita di potenza elettrica esterna per effettuare la misura. Lo svantaggio risiede invece nel produrre un campo elettrico proporzionale alla variazione di deformazione. Questo significa che gli accelerometri piezoelettrici possono misurare solo variazioni di accelerazione e non accelerazioni statiche (ossia a frequenza nulla).

Gli accelerometri piezoelettrici sono largamente riconosciuti come la miglior scelta per la misura assoluta delle vibrazioni. Il loro largo impiego è giustificato dalle seguenti proprietà che possono raggiungere:

- ampia banda dinamica: sono in grado di effettuare correttamente misure di accelerazione partendo da vibrazioni impercettibili (bassi g) fino ad arrivare ad alti shock (alti g).
- eccellente linearità su tutta la banda dinamica
- ampia banda frequenziale: sono in grado di effettuare misure di accelerazione ad elevata frequenza.
- trasduttori attivi: non richiedono alimentazione esterna.
- buona accuratezza anche in condizioni ambientali difficili
- estrema robustezza e ingombro contenuto.
- ampia banda di temperatura: sono in grado di effettuare misure anche con temperature maggiori di 120°C.

Da un punto di vista costruttivo esistono due tipi di accelerometri piezoelettrici:

- funzionanti a compressione
- funzionanti a taglio

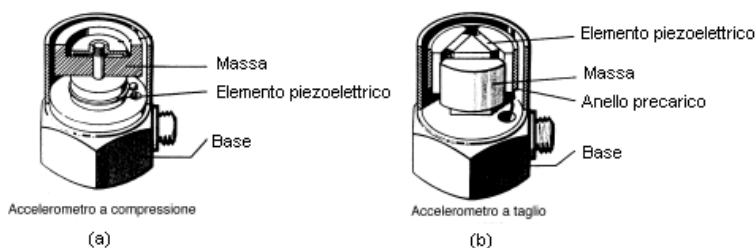


Figura 3. Accelerometri piezoelettrici a compressione (a) e a taglio (b).

Il primo tipo è illustrato in Fig. (3a). Il sistema massa-elemento piezoelettrico è fissato tramite un perno cilindrico centrale alla base dell'accelerometro. Ogni deformazione della base, dovuta a flessione o variazioni termiche, induce sollecitazioni nell'elemento piezoelettrico introducendo un errore nel segnale in uscita.

Per evitare questo fenomeno di disturbo si adotta la versione a taglio mostrata in Fig. (3b). In questo caso la massa sismica è fissata all'elemento sensibile in modo da esercitare un carico di taglio su di esso. In figura tre masse e tre cristalli piezoelettrici sono montati secondo le facce di un prisma a sezione triangolare ad un supporto centrale. Queste masse sono mantenute in posizione da un anello che esercita una azione di precarico che assicura una elevata linearità.

La carica che si sviluppa ai capi di un materiale piezoelettrico sottoposto a deformazione è proporzionale alla forza meccanica applicata:

$$P = d\sigma \quad (10)$$

in cui P è la densità di carica [C/m^2] e σ è la forza applicata per unità di area [N/m^2]. Il coefficiente di proporzionalità d è chiamato costante piezoelettrica [C/N].

Osservando la Fig. (4) supponiamo di sottoporre un materiale piezoelettrico presente in un accelerometro ad una forza F_x . Questa forza sia parallela all'asse X e agente sulle due facce di area A (elettrodi su cui si accumula la carica quando il materiale subisce la deformazione).

La carica che si accumula sugli elettrodi è:

$$q = AP = Ad\sigma = Ad \frac{F_x}{A} = dF_x \quad (11)$$

La carica è proporzionale alla forza applicata perciò, in accordo con la legge di Newton $F = ma$, è proporzionale anche alla massa e all'accelerazione. Questo dimostra come la sensibilità dell'accelerometro diminuisca al diminuire della massa.

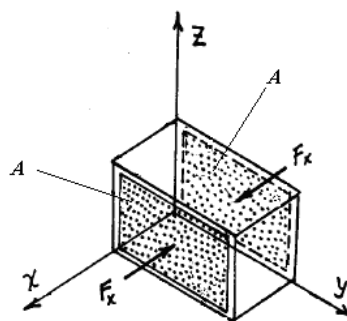


Figura 4. Elemento piezoelettrico sottoposto a deformazione.

Gli accelerometri piezoelettrici possono essere classificati in due categorie [5]:

- Accelerometri di carica (Charge Mode Accelerometers)
- Accelerometri IEPE (Internal Electronic Piezo-Electric)

I primi contengono solamente l'elemento sensibile senza elettronica

I secondi contengono internamente, oltre all'elemento sensibile, circuiti elettrici per il condizionamento del segnale.

Gli accelerometri piezoelettrici di carica generano segnali di uscita in termini di carica dell'ordine dei *picocoulombs*. Questi segnali sono caratterizzati da elevata impedenza e perciò il rischio di corruzione della loro integrità da parte dell'ambiente esterno non è da sottovalutare. Per effettuare misure corrette è necessario trasformarli in segnali in tensione a bassa impedenza.

Per operare questa trasformazione si utilizza uno schema circuitale noto come amplificatore di carica (Fig. (5)) [6]. Questo schema permette di ottenere una sensibilità dello strumento indipendente dalla capacità dei cavi di collegamento.

La capacità del sensore piezoelettrico è indicata con C . La capacità del filo che collega il sensore all'amplificatore di carica è indicato con C_c . La resistenza e la capacità di retroazione sono indicate rispettivamente con R_f e C_f .

Si supponga per semplicità che l'amplificatore operazionale sia ideale. In questo caso il nodo A è a massa e questo implica che la tensione sulla C_c vale zero. All'ingresso invertente non entra corrente perciò la legge di Kirchhoff delle correnti applicata al nodo A diventa:

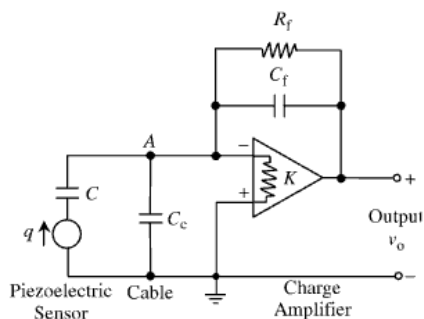


Figura 5. Amplificatore di carica.

$$\frac{dq}{dt} + \frac{v_0}{R_f} + C_f \frac{dv_0}{dt} = 0 \quad (12)$$

dove q è la carica generata dall'accelerometro e v_0 è la tensione di uscita.

Operando la trasformata di Laplace l'equazione può essere riscritta come funzione di trasferimento del sistema:

$$\frac{V_o(S)}{Q(s)} = -\frac{R_f S}{R_f C_f S + 1} \quad (13)$$

La formula appena scritta indica una chiara dipendenza tra la tensione di uscita e la carica in ingresso. Ricordando che $S = j\omega$ si nota che per $\omega = 0$ (frequenza zero) l'uscita assume valore nullo. In pratica un accelerometro piezoelettrico non può essere usato per misurare segnali costanti. Ad alte frequenze invece la funzione di trasferimento diventa:

$$\frac{V_o(S)}{Q(s)} = -\frac{1}{C_f} \quad (14)$$

La tensione di uscita dipende solo dalla carica e dalla capacità di retroazione.

Gli accelerometri dotati di una elettronica interna di condizionamento del segnale in grado di convertire il segnale in carica ad alta impedenza (generato dall'elemento sensibile) in un segnale in tensione a bassa impedenza, prendono il nome di accelerometri IEPE. Sebbene il segnale così ottenuto permetta di superare ostacoli legati alla catena di misura, ad esempio essere trasmesso attraverso lunghi fili conduttori con una minima degradazione, permangono alcuni limiti legati alla presenza dell'elettronica interna, ad esempio non possono essere impiegati sopra una certa temperatura di lavoro.

L'elettronica all'interno degli accelerometri IEPE richiede di essere alimentata tramite un generatore di tensione DC regolato a corrente costante.

3. PROVE SPERIMENTALI

In questo capitolo si riportano le prove sperimentali effettuate presso il laboratorio di Dinamica delle Macchine "F.Lombardini" del Dipartimento di Scienze e Metodi dell'Ingegneria (DISMI) dell'Università di Modena e Reggio Emilia.

Si sono individuati e confrontati fra loro otto diversi accelerometri di cui quattro di tipo MEMS capacitivo e quattro di tipo piezoelettrico. Per ciascuna tipologia si sono scelti un accelerometro triassiale e tre di tipo monoassiale.

In Tabella (1) riportano l'elenco degli accelerometri utilizzati con l'indicazione dei rispettivi produttori

La scelta degli accelerometri MEMS è stata effettuata tra i modelli che avessero un fattore di forma di tipo SOIC che ha permesso una più facile saldatura tra il componente e la scheda di condizionamento del segnale.

Tabella 1. Accoppiamento tra accelerometri MEMS e piezoelettrici.

Tipologia	Modello	Produttore	Assi di misura	Tipologia	Modello	Produttore
MEMS	MMA1211EG	Freescale	1	PIEZO	KS94B100	XXX
	MMA1200EG		1		352C41	PCB
	MMA1270EG		1		333B33	
	LIS3L02AS4	Analog Device	3		356A01	

Per ciascun accelerometro MEMS è stato progettato e realizzato un circuito per il condizionamento e una più agevole acquisizione del segnale. Inoltre su ciascuna scheda è stato ricavato uno spazio in cui poter incollare il corrispondente accelerometro piezoelettrico come mostrato in Fig. (6).

Operativamente le prove si sono svolte utilizzando uno shaker Data Physics Signal Force, capace di raggiungere i 3 kHz di vibrazione, per poter sollecitare in maniera opportuna gli accelerometri. Per garantire le stesse condizioni di sollecitazione a tutti gli accelerometri si è utilizzata una piastra metallica, avvitata all'expander dello shaker e sulla quale sono stati incollati i sensori, mediante colla a base di cianoacrilato.

La Tabella (2) riporta anche gli accoppiamenti tra le due tipologie di accelerometri che sono stati imposti per la comparazione durante le prove sperimentali. Ad esempio l'accelerometro MEMS MMA1211EG è stato affiancato all'accelerometro piezoelettrico KS94B100, ecc...

Le prove sono state suddivise in:

- prove statiche (MEMS)
- prove dinamiche (MEMS, PIEZO)

Le prove statiche sono state effettuate solo sugli accelerometri MEMS in quanto i piezoelettrici non sono in grado, come descritto nel capitolo precedente, di misurare una componente statica.

Le prove statiche sono di fondamentale importanza per definire due caratteristiche dei sensori che descriveremo in seguito: lo *zero-g level* e la *sensibilità*.

Tabella 2. Accoppiamento tra accelerometri MEMS e piezoelettrici.

MEMS	Range g	Sensibilità (mV/g)	Assi di misura	Piezoelettrico	Range g	Sensibilità (mV/g)
MMA1200EG	±250g	8	1	352C41	±500g	10
MMA1211EG	±150g	13,33	1	KS94B.100	±60g	100,9
MMA1270EG	±2,5g	750	1	333B33	±50g	100
LIS3L02AS4	±2g	660	3	356A01	±1000g	5

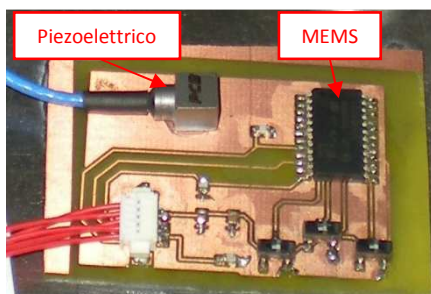


Figura 6. Accelerometri sullo shaker.

Prove Statiche

L'esigenza di eseguire delle prove statiche è nata dalla considerazione che gli accelerometri MEMS hanno valori di *sensibilità* e *zero-g level* definiti all'interno di una banda (stabilita da catalogo). Le motivazioni sono legate ai quantitativi di accelerometri prodotti. Infatti la produzione in larga scala dei MEMS ha probabilmente reso più conveniente definire una banda di conformità dei sensori piuttosto che realizzare una taratura di ciascuno di essi. Per una corretta lettura della misura è quindi necessario stabilire per ogni dispositivo sensibilità e zero-g.

La sensibilità rappresenta il coefficiente di proporzionalità che permette di tradurre la misura di tensione fatta ai capi del sensore in unità di accelerazione, normalmente in *g*.

Lo *zero-g level* è la tensione misurata ai capi dei sensori quando su di esso agisce un'accelerazione pari a zero *g*.

Per sottolineare l'importanza della definizione preliminare di questi due parametri, riportiamo per l'accelerometro MMA1200, una tabella (Tab. (3)) in cui le colonne riportano i valori minimi, tipici e massimi della sensibilità come da catalogo, e le righe riportano i valori minimi, tipici e massimi dello zero-g level.

Tabella 3. Caratterizzazione dei MEMS.

MMA 1200 (+/-250g) Accelerazione +1g			Sensibilità (mV/g)		
			Min	Tip	Max
			7,6	8	8,4
Zero-g (V)	Min	2,35	-17,8g	-17,75g	-17,7g
	Tip	2,5	0,95g	1g	1,05g
	Max	2,65	19,7g	19,75g	19,8g

Se si misurasse un'accelerazione di 1 g dando per scontato che entrambi i parametri abbiano i valori tipici, la Tab. (3) mostra le corrispondenti misurazioni nel caso in cui i valori reali dovessero spostarsi verso gli estremi degli intervalli ammissibili.

Ad esempio se l'accelerometro acquistato avesse un valore di zero-g di 2,65 V anziché 2,5 V (valore tipico), ipotizzando che la sensibilità reale sia coincidente con quella tipica, la lettura del sensore porterebbe a leggere un'accelerazione di 19,75 g anziché di 1 g!

Naturalmente l'effetto di una non corretta stima dello zero-g è più gravosa poiché si tratta di una grandezza in Volt, mentre la sensibilità opera nel campo dei millivolt.

Calcolo di sensibilità e zero-g

Per calcolare con esattezza la sensibilità e lo zero-g degli accelerometri MEMS sono state eseguite le seguenti prove. La piastra metallica su cui giacevano le schede elettroniche è stata posizionata in modo che i MEMS fossero rivolti verso l'alto e verso il basso con acquisizione dei dati in entrambe le posizioni. In questo modo è stato possibile applicare ai sensori ± 1 g di accelerazione. L'acquisizione dei dati (in tensione) è durata alcuni secondi con un periodo di campionamento di 50 kHz; è stato quindi possibile fare una valutazione statistica calcolando, per ogni accelerometro, il valore medio delle misure effettuate.

Gli accelerometri piezoelettrici normalmente non necessitano di una prova di sensibilità, poiché per ognuno di essi viene rilasciato un foglio di calibrazione da parte dell'azienda produttrice.

Per calcolare la sensibilità (in mV/g) di ognuno dei MEMS lungo l'asse di misura si è fatto uso della seguente formula:

$$Sensibilità = \left(\frac{Val.medio_{maggior} - Val.medio_{minore}}{2} \right) \cdot 1000 \quad (15)$$

dove con $Val.medio_{maggior}$ e $Val.medio_{minore}$ si sono indicati i valori medi delle misure effettuate rispettivamente a $+1g$ e a $-1g$.

Per il calcolo dello zero-g la formula utilizzata è:

$$Zero-g = \left(\frac{Val.medio_{maggior} + Val.medio_{minore}}{2} \right) \quad (16)$$

Di seguito sono riassunti i valori ottenuti nelle Tab. (4-5).

L'ultima colonna della Tab. (5), denominata *zero-g offset*, riporta la differenza tra il valore nominale dello zero-g ed il valore calcolato.

E' stata effettuata un'ulteriore prova per confermare i valori di zero-g calcolati analiticamente con la Eq. (16).

La piastra con i sensori è stata posizionata in verticale. Per garantire con precisione la perpendicolarità rispetto al piano di lavoro si è utilizzata una squadretta calibrata. In questa posizione l'asse di misura dei MEMS formava un angolo di 90° gradi rispetto alla direzione della gravita e la misura risultate era zero g. Le misurazioni effettuate hanno confermato la validità dei valori precedentemente calcolati.

Tabella 4. Sensibilità degli accelerometri MEMS.

	Accelerometri MEMS	Sensibilità catalogo (mV/g)			Prova statica (accelerazione di gravità)
		Min	Tip	Max	Sensibilità (mV/g)
monoassiali asse z	MMA1200	7,6	8	8,4	7,5
	MMA1211	12,66	13,33	14	12,8
	MMA1270	712,5	750	787,5	771,9
triassiale (+/-2g)	LIS3L02AS4(z)	594	660	726	581,9

Tabella 5. Zero g e zero g offset degli accelerometri MEMS.

	Accelerometri MEMS	Zero-g catalogo (V)			Prova statica (accelerazione di gravità)	
		Min	Tip	Max	Zero g (V)	Zero g offset (mV)
monoassiali asse z	MMA1200	2,35	2,5	2,65	2,4973	2,7
	MMA1211	2,35	2,5	2,65	2,4864	13,6
	MMA1270	2,25	2,5	2,75	2,4421	57,9
triassiale (+/-2g)	LIS3L02AS4(z)	1,485	1,65	1,815	1,5279	122,1

Prove Dinamiche

Le prove dinamiche sono state eseguite per comprendere il comportamento dei MEMS a fronte di ingressi differenti valutandone qualitativamente e quantitativamente le prestazioni rispetto ai rispettivi sensori piezoelettrici.

E' stato utilizzato uno shaker Data Physics Signal Force e un sistema di acquisizione National Instruments costituito da una base NI-cACQ 9172 e dai moduli NI-9239 (per l'acquisizione dei MEMS) e NI-9234 (per l'acquisizione dei piezoelettrici). Nel corso delle prove è stata sollecitata ogni coppia di accelerometri con un ingresso sinusoidale a diversa frequenza. Le prove sono state articolate nel seguente modo:

- MMA1200-352C41: accelerazione sinusoidale 30g a frequenze di 100Hz, 200Hz, 300Hz, 400Hz. La prova a 100 Hz è stata ripetuta 3 volte intervallata da un ingresso random dello shaker.

Tabella 6. Deviazione standard media per ogni accelerometro alle diverse frequenze di esercizio.

	Frequenza (Hz)	Deviazione standard media (g)		Frequenza (Hz)	Deviazione standard media (g)
LIS3L02AS4 2g sin	10	0,00753	MMA1270 2,5g sin	10	0,00640
	10_random 1	0.00710		10_random1	0,00690
	10_random 2	0,00720		10_random1	0,00650
	50	0,01950		20	0,00980
	100	0,03650		30	0,01340
	400	0,13100		40	0,01540
356A01 2g sin	10	0,06375	333B33 2,5g sin	10	0,05010
	10_random 1	0,06370		10_random1	0,05030
	10_random 2	0,06300		10_random1	0,05060
	50	0,14040		20	0,05240
	100	0,20650		30	0,06790
	400	0,21150		40	0,08880
	Frequenza (Hz)	Deviazione standard media (g)		Frequenza (Hz)	Deviazione standard media (g)
MMA1211 30g sin	100	0,58330	MMA1200 30g sin	100	0,61190
	100_random1	0,58530		100_random1	0,60920
	100_random2	0,58650		100_random2	0,61090
	200	1,04400		200	1,12830
	300	1,43050		300	1,52320
	400	1,57620		400	1,68930
KS94B 30g sin	100	2,63360	352C41 30g sin	100	3,33080
	100_random1	3,42220		100_random1	3,77490
	100_random2	3,51760		100_random2	3,74020
	200	4,92520		200	5,48690
	300	6,11790		300	6,74960
	400	7,37150		400	8,75810

- MMA1211-KS94B: accelerazione sinusoidale 30g a frequenze di 100Hz, 200Hz, 300Hz, 400Hz. La prova a 100 Hz è stata ripetuta 3 volte intervallata da un ingresso random dello shaker.

- MMA1270-333B33: accelerazione sinusoidale 2,5g a frequenze di 10Hz, 20Hz, 30Hz, 40Hz. La prova a 10 Hz è stata ripetuta 3 volte intervallata da un ingresso random dello shaker.
- LIS3L02AS4-356A01: accelerazione sinusoidale 2g a frequenze di 10Hz, 50Hz, 100Hz, 400Hz. La prova a 10 Hz è stata ripetuta 3 volte intervallata da un ingresso random dello shaker.

Si noti che a seconda dei range di accelerazione degli strumenti è stata cambiata l'accelerazione in ingresso. Sono state fornite accelerazioni elevate agli accelerometri in grado di riceverle, nel limite fisico dello shaker. Allo stesso modo la frequenza di esercizio è stata variata tenendo conto della banda di linearità di ogni dispositivo.

Per valutare la precisione dei sensori è stata fornita per tre volte la stessa eccitazione intervallata da segnali random.

Discussione

In considerazione della precedente Tab. (6) è possibile affermare che:

- La deviazione standard dei MEMS è sempre minore di quella dei piezoelettrici.
- La deviazione standard aumenta all'aumentare della frequenza di lavoro.
- La deviazione standard assume valori simili quando si effettuano prove ripetute ad una stessa frequenza intervallate da un ingresso random. Questo indica che lo strumento mantiene la stessa precisione ad ogni acquisizione.

Nella maggior parte delle prove la forma sinusoidale in ingresso è stata misurata correttamente sia dagli accelerometri MEMS che dai piezoelettrici. In Fig. (7) sono messi a confronto come esempio gli andamenti sinusoidali medi degli accelerometri MMA1211 e KS94B.

Si noti che il piezoelettrico ha una curva più frastagliata rispetto a quella del MEMS.: in condizioni di prova si sommano alla sinusoide in ingresso delle accelerazioni spurie ad alta frequenza – probabilmente indotte dall'apparato sperimentale – che vengono rilevate solo dal piezoelettrico in quanto il sensore MEMS è dotato di un filtro passa-basso RC nel circuito di condizionamento del segnale.

La deviazione standard media è per tutti i dispositivi MEMS minore rispetto alla deviazione standard media misurata nei risultati dei piezoelettrici. Probabilmente questo è imputabile al fatto che il piezoelettrico ha un segnale di misura molto frastagliato come riportato in Fig. (8).

Le differenze riportate per i segnali nel dominio del tempo si confermano confrontando i corrispondenti spettri in frequenza. Le Figure (9 – 11) evidenziano come lo spettro dei MEMS mantenga un livello di rumore più elevato in tutto il campo dello spettro.

Tuttavia vi sono state prove che hanno mostrato un rumore maggiore nei piezoelettrici. E' il caso rappresentato in Fig. (11). Entrambi i dispositivi hanno l'armonica fondamentale a 300 Hz ma il piezoelettrico risente di maggiore rumore. Il motivo potrebbe risiedere nella maggiore capacità di misurare vibrazioni ad alta frequenza da parte del piezoelettrico.

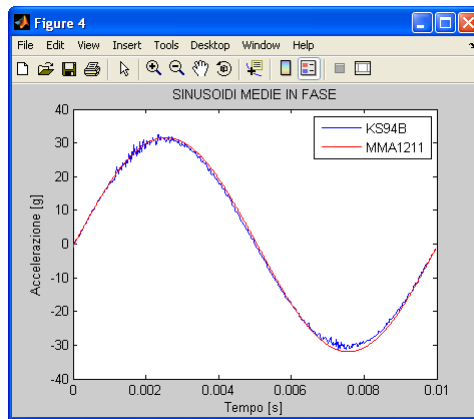


Figura 7. Segnali sinusoidali acquisiti da MMA1211 e KS94B a 100 Hz.

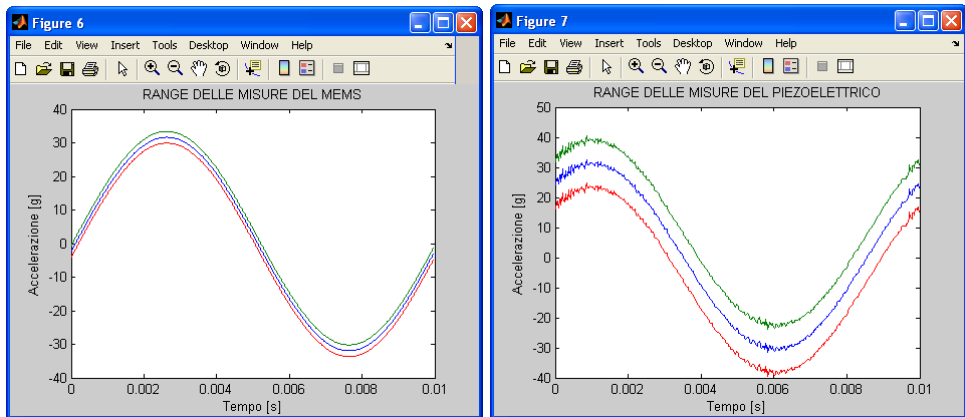


Figura 8. Banda di ampiezza che include il 99% delle misure per i dispositivi MMA1211 e KS94B a 100Hz.

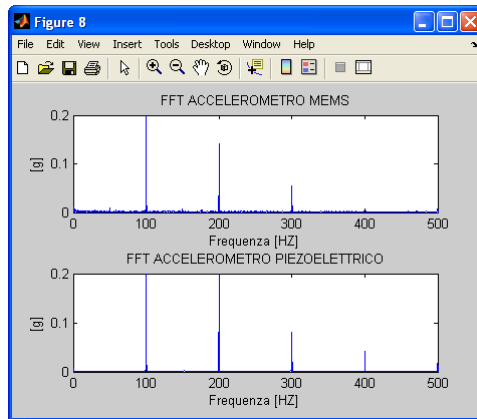


Figura 9. Ingrandimento dello spettro per gli accelerometri MMA1211 e KS94B a 100Hz.

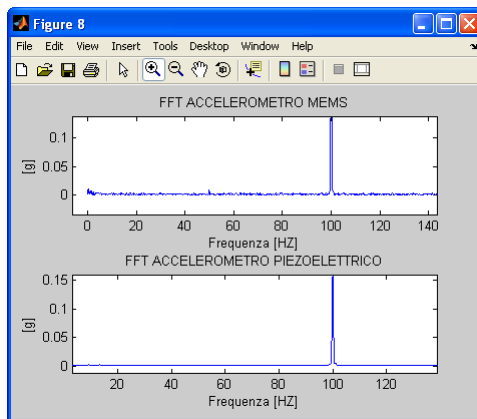


Figura 10. Rumore nell'intorno della fondamentale per i dispositivi MMA1211 e KS94B.

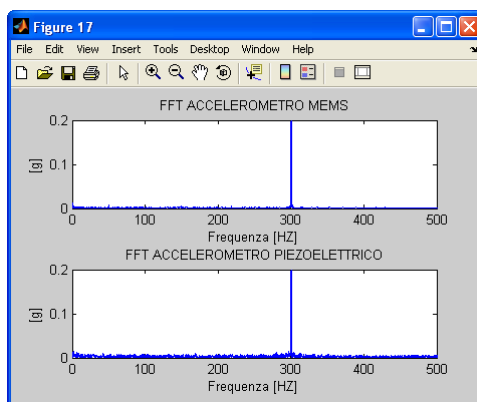


Figura 11. Rumore per MMA1211 e KS94B a 300Hz.

4. APPLICAZIONE PRATICA

In questa parte si riportano i risultati ottenuti utilizzando gli accelerometri per l'analisi modale di una lamina metallica.

Per lo sviluppo di questa applicazione la piastra che ospitava gli accelerometri è stata fissata all'estremità del piatto dello shaker assumendo la configurazione di una trave a sbalzo (Fig. (12)). Successivamente la parte sporgente della piastra è stata sollecitata utilizzando un martello strumentato e la vibrazione è stata acquisita dai differenti accelerometri.

Per ogni accelerometro sono state calcolate la funzione di risposta in frequenza (FRF) e la funzione densità spettrale incrociata adimensionalizzata.

La FRF è la funzione di trasferimento che correla l'ingresso imposto (impulso) all'uscita (accelerazione misurata).



Figura 12. Piastra in configurazione a sbalzo.

La funzione densità spettrale incrociata adimensionlizzata è un parametro che stima il grado di coerenza tra ingresso ed uscita del sistema, essa restituisce per ciascuna linea spettrale, un valore compreso tra 0 e 1 , che corrispondono rispettivamente alla minima ed alla massima correlazione tra eccitazione e risposta.

Di seguito (Fig. (13-20)) sono riportati i grafici di coerenza e FRF per ogni accelerometro

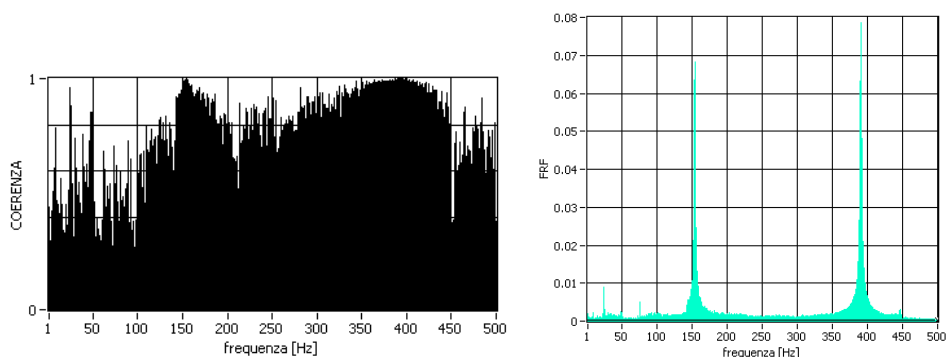


Figura 13. Coerenza e FRF per MMA1200.

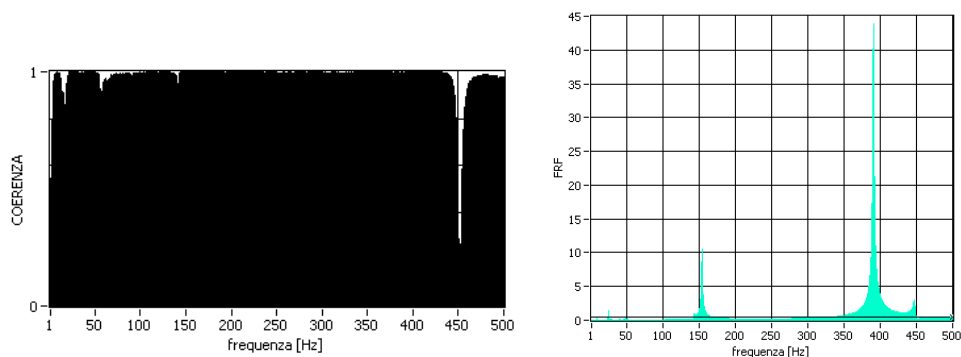


Figura 14. Coerenza e FRF per 352C41

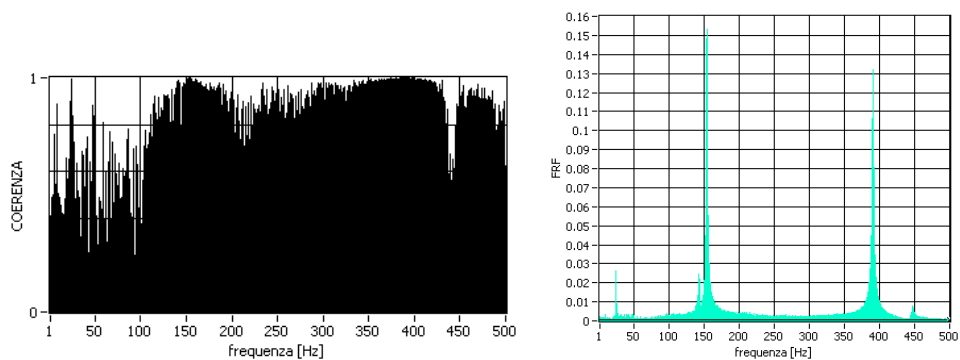


Figura 15. Coerenza e FRF per MMA1211.

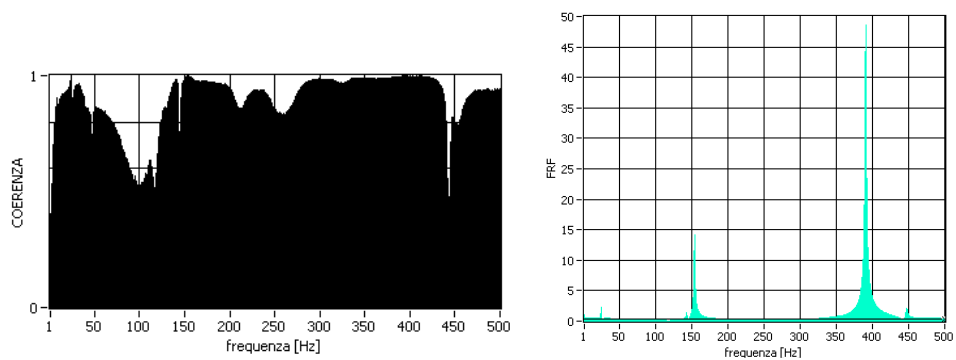


Figura 16. Coerenza e FRF per KS94B.

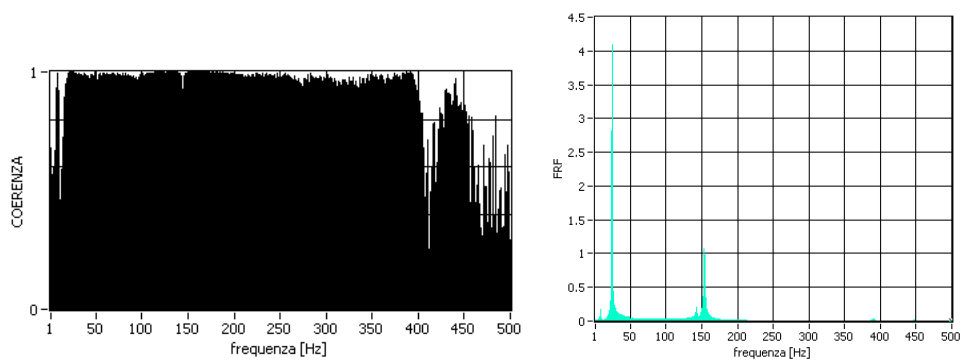


Figura 17. Coerenza e FRF per MMA1270.

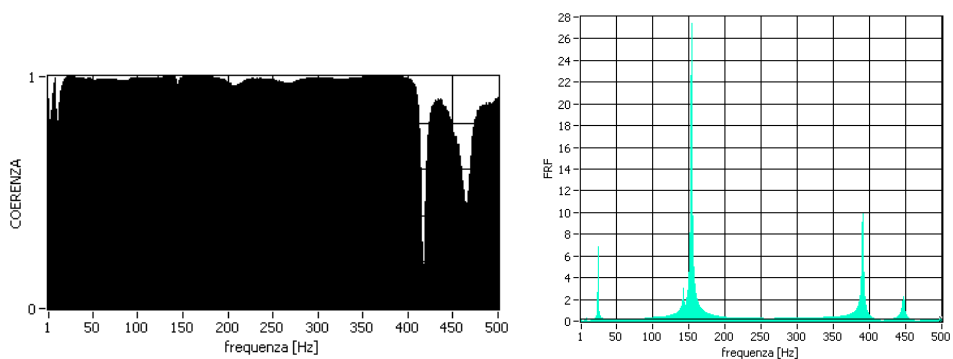


Figura 18. Coerenza e FRF per 333B33.

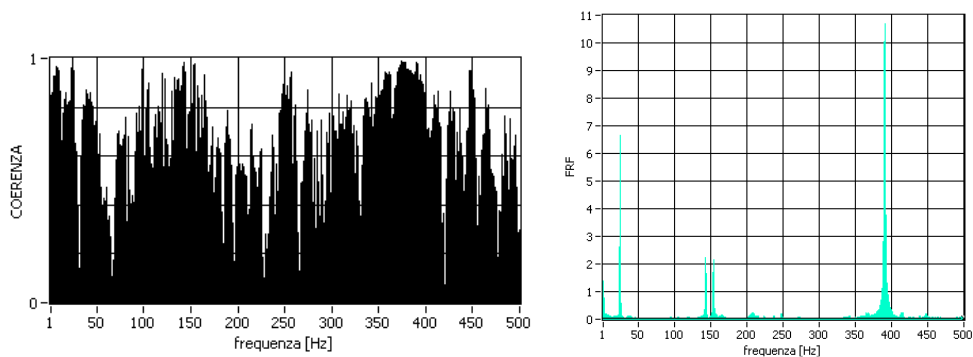


Figura 19. Coerenza e FRF per LIS3L02AS4.

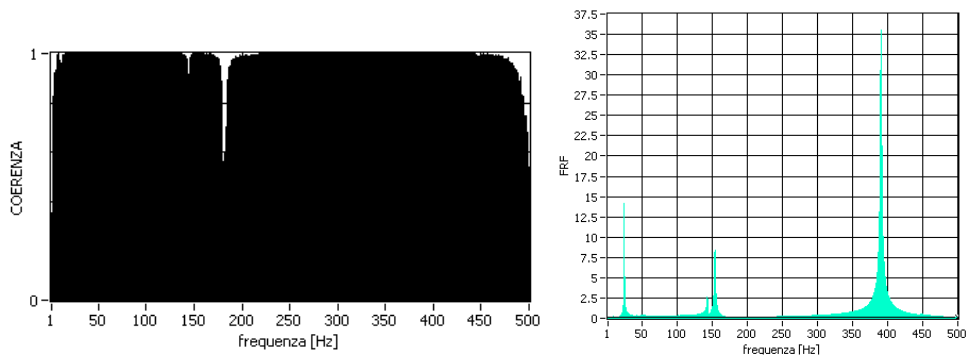


Figura 20. Coerenza e FRF per 356A01.

Confrontando l'accelerometro MMA1200 con il piezoelettrico 352C41 si nota che entrambi restituiscono i medesimi auto valori. I picchi prossimi ai 150 Hz e 400 Hz vengono valutati da entrambi gli accelerometri come realmente associati all'ingresso

fornito, in quanto la coerenza corrispondente è prossima a uno. A differenza del sensore piezoelettrico la coerenza valutata per il MEMS è inferiore a uno per le basse frequenze.

Quanto detto è valido anche per il MEMS MMA1211 e il relativo accelerometro piezoelettrico KS94B.

Il piezoelettrico 333B33, a differenza del MEMS MMA1270, riconosce alle alte frequenze una risonanza a cui corrisponde una coerenza prossima all'unità. La funzione di coerenza dei due sensori è comunque simile.

La coerenza del MEMS triassiale è molto differente se confrontata con il relativo piezoelettrico triassiale. I principali picchi vengono riconosciuti da entrambi i trasduttori e per questi il valore di coerenza è comunque simile. Nell'intorno dei 150 Hz il MEMS misura due picchi molto vicini tra loro in modo equivalente al piezoelettrico. Inoltre localizza picchi di ampiezza ridotta lungo tutta la banda in frequenza non individuati dal piezoelettrico.

La FRF è definita anche per valori di frequenza maggiori della banda degli accelerometri (il caso più evidente è il dispositivo MMA1270 che ha una banda di 50Hz).

I picchi che si trovano in questa banda di frequenze forniscono solamente informazioni qualitative sulla risposta del sistema. Infatti al di fuori di tale banda gli accelerometri non hanno un comportamento lineare.

5. CONCLUSIONI

In questo articolo vengono messi a confronto accelerometri MEMS capacitivi e accelerometri piezoelettrici. E' noto che i sensori piezoelettrici rivestono un ruolo di eccellenza in ambito industriale, ove diviene strategico decrementare i costi a parità di prestazioni.

Lo scopo della ricerca è stato la valutazione delle effettive potenzialità dei sensori a basso costo MEMS paragonate a quelle note dei piezoelettrici.

Durante lo svolgimento del lavoro i dati acquisiti dagli accelerometri piezoelettrici sono stati presi come riferimento per valutare le prestazioni dei MEMS.

Per poter utilizzare correttamente entrambi i tipi di accelerometri è stato necessario valutare due parametri caratterizzanti il sensore: la sensibilità e lo zero-g level. Questi sono ben definiti nel caso di accelerometri piezoelettrici mentre si trovano all'interno di una banda per i MEMS. Per questa tipologia di accelerometri si è verificato che l'errore di misura che si commette quando non si conosce lo zero-g è maggiore dell'errore che si commette quando non si conosce la sensibilità del sensore, ed è stato dimostrato sperimentalmente come l'utente non possa esimersi dalla loro caratterizzazione prima di poter effettuare delle misure. Queste misure sono meno rilevanti per i piezoelettrici i quali hanno un valore zero-g prossimo allo zero e la cui sensibilità è normalmente riportata sul foglio di calibrazione fornito dal produttore.

Un'importante indicazione a favore della tecnologia MEMS è stata ottenuta dalle prove dinamiche eseguite mediante l'impiego di uno shaker. In particolare è stata eseguita un'analisi modale di una piastra metallica sulla quale sono stati applicati gli accelerometri. I dati di accelerazione sono stati elaborati ottenendo la funzione di coerenza e la FRF. I risultati ottenuti mostrano come i MEMS siano in grado di riconoscere nella maggior parte dei casi le risonanze del sistema con un indice di certezza prossimo a quello dei piezoelettrici. La prova permette di valutare la tecnologia MEMS come una valida alternativa alla tecnologia piezoelettrica.

Ringraziamenti

Si ringrazia l'ing. Guido Pedrini per il contributo dato durante l'attività sperimentale e per lo scambio di idee e suggerimenti.

Questo lavoro è supportato dal tecnopolo InterMech-MoRE.

BIBLIOGRAFIA

- [1] Kaajakari, V., 2009, *Practical MEMS*, Small Gear Publishing, pp. 1-5.
- [2] Maluf N., 2004, *An Introduction to Microelectromechanical Systems Engineering*, 2nd ed, Artech House Inc., Norwood, pp. 135-137.
- [3] Kaajakari, V., 2009, *Practical MEMS*, Small Gear Publishing, pp. 379-382.
- [4] Bustillo, J.M., Howe, R.T., Muller, R.S., "Surface micromachining for microelectromechanical systems", *Proceedings of the IEEE*, vol. 86, no. 8, pp. 1552-1574.
- [5] Gardner, J.W., Varadan, V.K., and Awadelkarim, O.O., 2001, *Microsensor, MEMS and Smart Devices*, John Wiley & Sons Ltd, Chichester, pp.145-149.
- [6] Kovacs, G.T.A., Maluf, N.I., and Petersen, K.E., 1998, "Bulk micromachining of silicon", *Proceedings of the IEEE*, vol. 86, no. 8, pp. 1536-1551.

INVESTIGATION AND PREDICTION OF CHARGE WELDS IN ALUMINUM EXTRUDED PROFILES

Antonio Segatori

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: antonio.segatori2@unibo.it*

Barbara Reggiani

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: barbara.reggiani@mail.ing.unibo.it*

Lorenzo Donati

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail l.donati@unibo.it*

Floriano Bagagli

*Alutitan,
Repubblica di San Marino,
E-mail: florianobagagli@alutitan.com*

Luca Tomesani

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: luca.tomesani@unibo.it*

Abstract. *The direct aluminum extrusion can be considered as a continue process although the billets are discretely loaded into the press. Indeed, inside the die, the high hydrostatic pressure field produces the joining of the material into a single bar so as to consider the production as continue. However the mechanical proprieties of the profile are lower in the welded region compared to the adjacent ones thus requiring the scrapping of the welded segment. The extension of the segment is primarily affected by the interaction of seam and charge welding phenomena. The billet skin inclusion on the profile has to be accounted for as an additional contaminating effect . The computation and localization of the segment to be discarded is a complex matter that actually involves approximate empirical relations or labor intensive analyses. Aim of the paper is to investigate the interaction between the different welding phenomena inside an industrial multi-profile die, in order to determine the exact position and the minimal profile length to be discarded. The extrusion of four hollow profiles made of AA6060 alloy through a multi-hole die has been initially monitored in an industrial press: process loads, temperatures and speeds were accurately recorded during the process. The four profiles were sectioned starting from the 'stop mark' then grinded and etched in order to investigate the location and the dimension of the zone to be*

ridged. Finite element simulations of the process were then carried out by means of the commercial code Altair® HyperXtrude®. Numerical predictions were then compared to experimental data with a particular focus on the computation of charge weld evolution. The simulated data accurately predicts the evolution of the phenomenon and the dimension of the profiles to be discarded. As conclusion, the codes resulted as a reliable tool for the localization and determination of the scraps, in particular if new die design strategies has to be evaluated also in term of reduction of the zone extension.

Keywords: *extrusion, charge weld, seam weld, finite element modelling*

1. INTRODUCTION

Aluminum extruded profiles are today extensively used in several type of applications like buildings and construction, engineering, transport, up to airplane primary structures. The proprieties and quality of the profiles are strictly related to their final application: if for aesthetical applications (i.e. windows frames or interior design structures) the surface aspect is of primary importance, in more severe loaded conditions (like automotive or aerospace) the mechanical proprieties of the profiles remains of critical importance. Parson, Hankin and Bryant [1] classified and described the most common extrusion defects: pick-up, blistering, streaks and back end defects play a primary role on the surface aspects while in term of mechanical proprieties the most critical problems are incorrect heat treatment, seam welds failure and presence of charge welds inside the profile. Such defects often mutually interact but the studies available in literature usually analyze each one singularly.

Back end defect. It is also often called contamination by billet skin. In the direct hot extrusion of aluminum profiles a complex material flow is generated in the container and in the die in relation to the high friction conditions generated at the aluminum-steel interface. The friction produces an imbalanced material flow that on one side positively affects the process (reduction of contamination by billet skin) but on the other side generates dead metal zones and complex material flow inside the die. The skin of the billet usually is different from the inner material; this difference can be the results of an altered chemical composition (as consequence of cooling from casting) or of a contamination due to impurities (oxides, dust, oil, etc) during billet shipping, pre-heating and loading in the press. The key parameter used for controlling the billet skin contamination is the dimension of the discarded butt: a longer butt allows holding more billet skin and consequently it guarantees a contaminated free profile, while shorter butts represent a profit because less material is discarded. Figure 1 shows the evolution of the billet skin during a typical un-lubricated hot extrusion processing of an aluminum alloys: in relation to the high friction factor with the container the billet skin is accumulated in proximity of the ram and, if an adequate butt length is used, it is entirely discarded (Fig. 1a), while, if a too short butt is used, the billet skin can flow inside the die up to the welding chambers and eventually also in the profile (Fig. 1b). The phenomenon has been extensively investigated in literature [2, 3, 4]. In order to understand if the butt length is adequate microstructure analysis of the profiles in proximity of the stop mark has to be done. When the ram is stopped the profile stick to the bearing zones thus creating a very evident mark on the profile surface called “stop mark”: if the billet skin overcome the stop mark the contaminated part of the profile has to be cut and discarded, while if the skin remains inside the die (eventually up to the welding chamber)

the profile is considered skin free. In the industrial environment this type of analysis is performed by cutting several slices of profile on the right and left side of the stop mark, grind and etch them in order to verify the complete absence of skin contamination.

Seam weld. When hollow profiles are required, the die is split into two parts, the mandrel and the die (Fig. 2). The mandrel realizes the shape of inner cavities while the die the outer shape. The billet is initially divided around the mandrel legs then it rejoins in the welding chamber thus realizing a solid state welding all along the profile length; such weld is called “seam or longitudinal weld”. The seam weld usually is marked by high mechanical proprieties because it is a solid state junction generated at high hydrostatic pressure under oxide-free conditions.

Charge weld. The material flow becomes even more complex when multiple billets are consecutively loaded into the press, as usually happens during continuous extrusion. In these conditions, a weld is generated at the interface between the old and the new billet; it is the so called “charge or transverse weld” (Fig. 3). The charge weld is also generated at high hydrostatic pressure, but differently from seam welds, it usually carries contamination of oxides, dust or lubricant consequently the loading into the press. As consequence, the charge weld has to be discarded from the final profile because it performs lower mechanical proprieties.

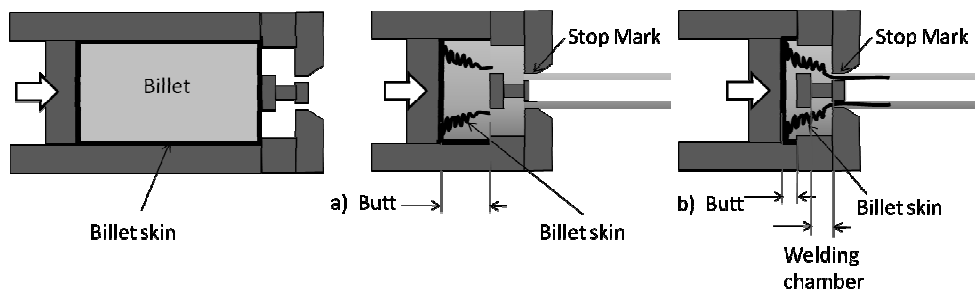


Figure 1. Billet skin contamination in the extrusion of a tube (schematic): a) long butt, b) short butt.

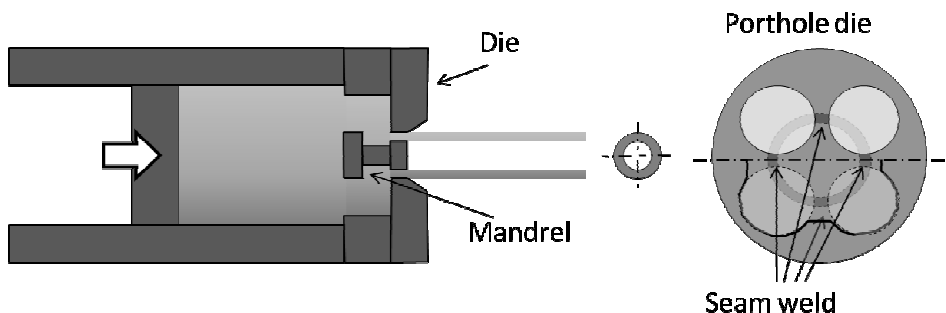


Figure 2. Schematic representation of mandrel die and seam weld localization.

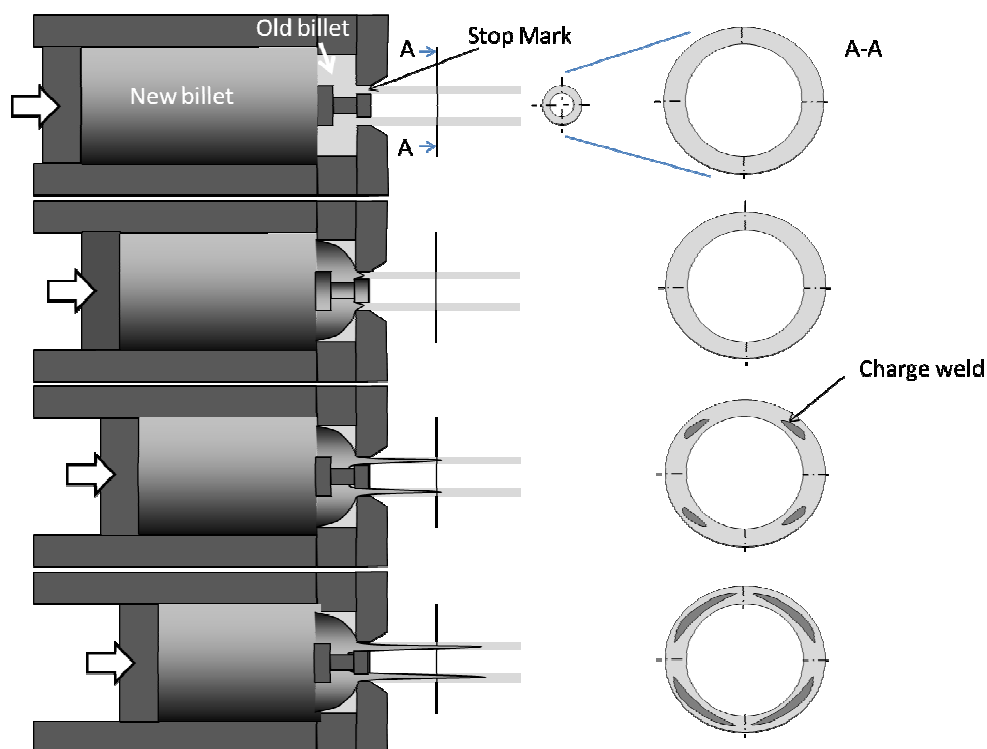


Figure 3. Sequence of a charge weld formation in a longitudinal and transverse views.

The three aforementioned phenomena have a deep interaction on the defect magnitude and extent. As a consequence, a wide segment of the final profiles has to be discarded due to lower mechanical proprieties. In industrial practice, the determination of the zone to be discarded is usually performed by means of rule of thumb or, in more critical applications, by labor intensive microstructure investigations at the beginning and end of each extruded lots. These results in a sensible process inefficiency since either large quantities of profile has to be discharged or great amount of time have to be spent for scrap minimization. In addition, it is often not taken into account the possibility to reduce this defect through a dedicated die design. Nowadays, these defects are generally separately described and analyzed in literature.

Finite element modelling (FEM) is a powerful tool currently used as support for many product and process optimization activities thus avoiding expensive and time-consuming trial-and-error experimental procedures. In the last decade, in a number of works in literature the extrusion process of aluminum profiles has been simulated by means of FE codes [5]. The common effort is to improve the computational tools in order to limit the role of expensive experimental campaign. In the extrusion context while the seam welds evolution has already been modelled by means of a number of numerical models [9, 10, 11, 12, 13], to the knowledge of the authors, no previous studies report the numerical validation of the charge weld evolution.

Aim of the present study was to investigate the interaction between the different welding phenomena inside an industrial multi-profile die, in order to determine the exact position and the minimal profile length to be discarded. Finite element simulations of the process were then carried out by means of the commercial code Altair® HyperXtrude® to evaluate the numerical code capabilities for a real industrial application. Numerical predictions were then compared to experimental data with a particular focus on the computation of charge weld evolution.

2. EXPERIMENTAL INVESTIGATION

Experimental set-up

The benchmark experiment was performed at the Alutitan s.p.a in San Marino, Italy, by means of a 1600 tons extrusion press. The profile was selected among a number of different designs available for the experiment on the basis of two main constraints. On one end, the presence of a thick, massive, part was required to ease the charge weld localization. On the other hand, the absence of too small details like tips, tongues, brush paths guaranteed the modellability of the profile with a reasonable number of finite elements and, at the same time, reduced the risk of a premature failure of the die. The final die consisted of 4 L-shaped holes to produce semi-hollow profiles as shown in Fig. 4. The minimum profile thickness was 1 mm, compatible with the requirement of an easy meshable component. The area of each profile was 135 mm² and the maximum bearing length 11 mm.

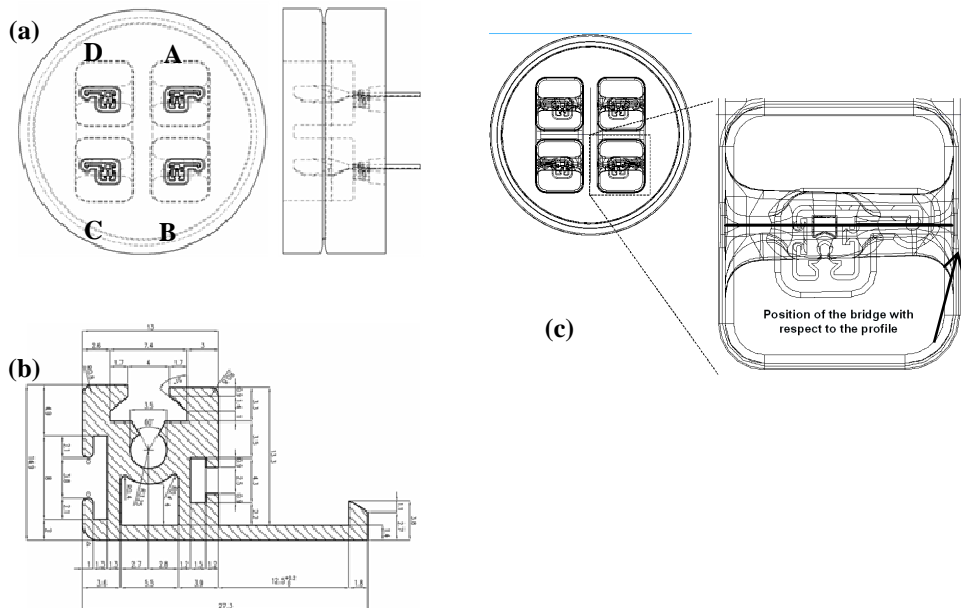


Figure 4. (a) The die used in the experimental design with the profile labelling, (b) shape and dimensions of the profile and (c) position of the bridge with respect to the profile geometry for the seam weld formation.

Table 1. Initial temperatures and geometries of the billets and tools involved in the experimental set-up.

	Temperature [°C]	Lenght [mm]	Diameter [mm]
Billets	478	500	178.0
Die	468		
Container	410	1150	186.0
Ram	450	150	185.5

The presence of multiple holes allowed to investigate, within the same experiment, the influence of even small differences in the aluminium flow stress on the charge weld evolution due to the asymmetric position of profiles A and D with respect to B and C. Four billets made of a AA6060 alloy were consecutively extruded during the experimental investigation. The temperature of the container was 410°C ; that of the ram 450°C. The die and the billet were preheated at 468 °C and 478 °C respectively. In Tab. 1 are reported the initial thermal and geometrical conditions of the experimental set-up.

For the first three billets, only the temperature of the profile A was recorded at different times by means of a thermocouple manually kept at 2.5 m distance from the die exit. Additionally, the temperature of the same profile was traced all over the extrusion time with a pirometer placed at 1 m from the die exit. Indeed, for the last billet the temperature of all the four profiles was acquired. The extrusion load, ram speed and profile velocity were registered as well.

The starting reference point was the stop mark, showed in Fig. 5b, a very evident mark on the profile surface due to the sticking of the material on the bearing zones when the ram is stopped. In Fig. 5 are also showed the four extruded profiles and the labelling method used to track the sections of the profiles.

The transition between the first two billets was chosen to explore the welding phenomena interaction in order to limit the interference of previous billet material deposits. The extruded profiles were then analyzed by cutting several slices spaced of 5 cm on the die side from the stop mark for welds investigation. Each specimen was then grinded and etched (HF 10%) to be observed at the optical microscope. For each slice, the percentage area of the new billet was computed by means of a dedicated tool included in the microscope software. The analysis was carried out till complete exhausting of charge welds. Analogous activity was performed on the other side of the profile from the stop mark to evaluate back end defects.

Experimental results

In Tab. 2 are reported the process parameters that were monitored during the experiment for the four extruded billets.

The lower ram speed for the recording point 1.1 was due to the acceleration gradient of the ram. Since the profiles were supported by a puller, the speed was stable and only the first recording was performed with free profile prior to tip cutting.

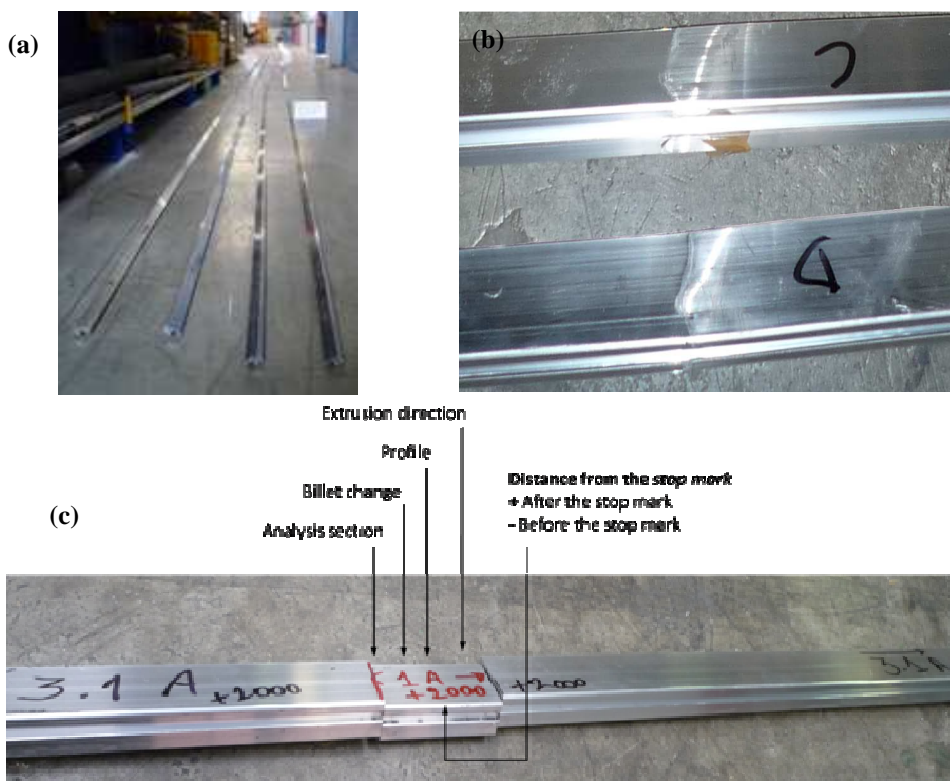


Figure 5. (a) The four extruded profiles, (b) the stop mark evidence in two of the extruded profiles and (c) the experimental sample labelling.

This resulted in a different profile speed only for the first recording even if with a higher ram speed. Similarly the extrusion load of the first billet was still influenced by the initial thermal transient when the die is at a lower temperature and therefore increase the flow stress of the material. The pyrometer temperature at recording point 1.2 was not considered as reliable and simply attributed to a disturbed recording.

At first it was investigated the trend and quality of the seam weld. For this activity the tips of each profile were sliced and analyzed from 50mm out from the tip. The welds were well evident has a dark line (Fig. 6). They were localized, as expected, close to the two bridges although not perfectly straight and a bit shifted toward the thinner part of the profile.

Figure 7 shows their position in respect to the bridges axes. This effect is most likely to be associated to a faster flow in the more massive part, also in good agreement with the results of the charge welds shown below.

Table 2. Process parameters monitored during the experimental campaign.

Billet	Recording point	Profile	Speed [mm/s]		Temperature [°C]*		Extrusion load [tonn]
			Ram	Profile	Thermocouple	Pyrometer	
1	1.1	A	5.3	440	530	529	1299
	1.2	A	6.3	340	560	666	1299
2	2.1	A	6.1	320	520	570	1183
	2.2	A	6.4	320	560	538	1023
3	3.1	A			560	524	
4	4.1	A	6.2	320	520	536	1183
	4.2	A	6.1	320	555	520	1048
	4.3	A			496		
		B			530		
		C			560		
		D			560		

The weld visibility, caused by impurities, ended within 150-250mm from the tip, after which no line or inhomogeneous grain structure were visible.

The investigations of back end defect did not reveal any contamination of the billet skin in the profile. Both the profile section before and after the stop mark did not show the typical variation in microstructure color (mostly lighter) associated with oxide or carbide contamination.

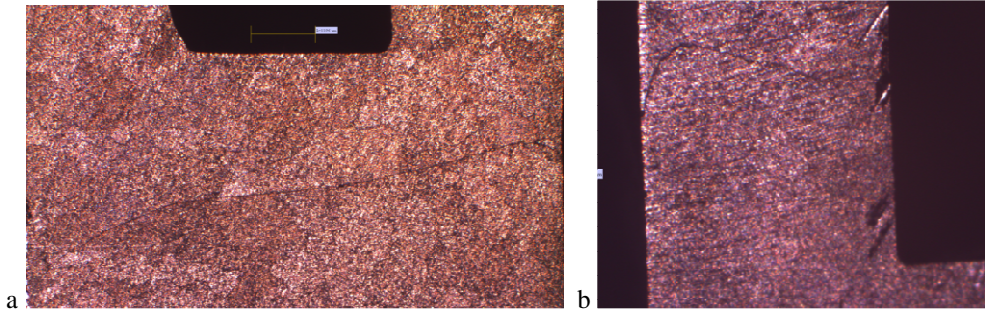


Figure 6. (a) Left seam weld for profile A at 200mm from tip – bright field, (b) right seam weld for profile A at 200mm from tip – bright field.

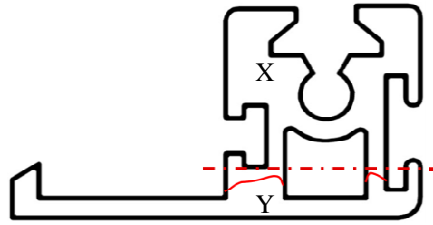


Figure 7. Seam weld location (solid line) compared to bridge axes (interrupted line) that divides the material into flows: X the massive part and Y the thinner one.

The analysis of the charge weld produced micrographs as the one shown in Fig. 8, were multiple micrographs have been composed to recreate a complete view of the weld over the profile section. Once the image was composed, it was then possible to carry out a measure of the area enclosed by the weld line and therefore the percentage of new material at a specified point of the profile, that is, the exhausting of the old billet material. Along with this activity the starting and ending point were looked for. With both this information it was possible to retrieve extent, position, and trend of charge weld for each of the four profiles.

Table 2 reports different section analyzed profile A with the weld line highlighted. Figure (a) well shows the starting point of the charge weld which appears always as a closed loop merging to surface. Following pictures in the table show the weld with the increase of distance. It is well evident in Fig. (e) as the new material has completely filled the upper part of the profile and the weld tend to arrange over the seam weld path. Less than 100mm further, Fig. (f), there is no trace of charge weld in the upper part of the profile -i.e. it has been completely filled with the new billet material- while the lower flux charge weld tip is visible in the thicker part of the lower section. Further images show the filling of the lower part. Two results are of immediate evidence: first, the lower -i.e. thinner part- flux presents a longer charge weld extent; second, the filling of the two part of the profile by the new billet material is completely disjointed, that is, there is no overlapping of charge weld for the two fluxes, thus increasing the final extent of the charge weld.

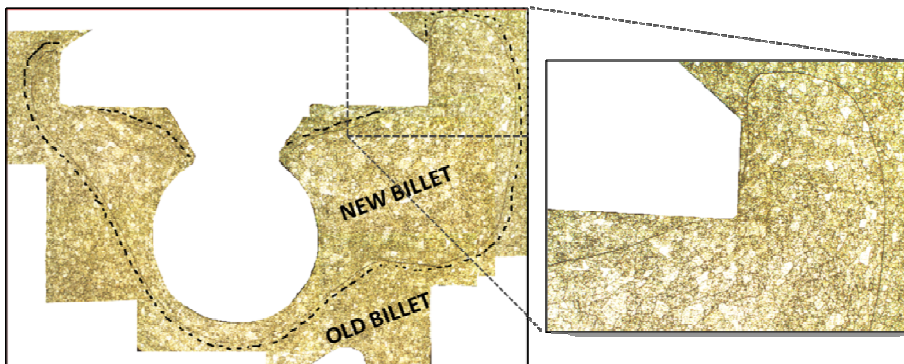


Figure 8. Composition of complete micrographical profile section

Table 2. Micrographics of the extruded profile with the charge weld mark.

(a) 565	(b) 600
(c) 700	(d) 750
(e) 950	(f) 1090
(g) 1150	(h) 1300
(i) 1600	(l) 1750
	(m) 1900

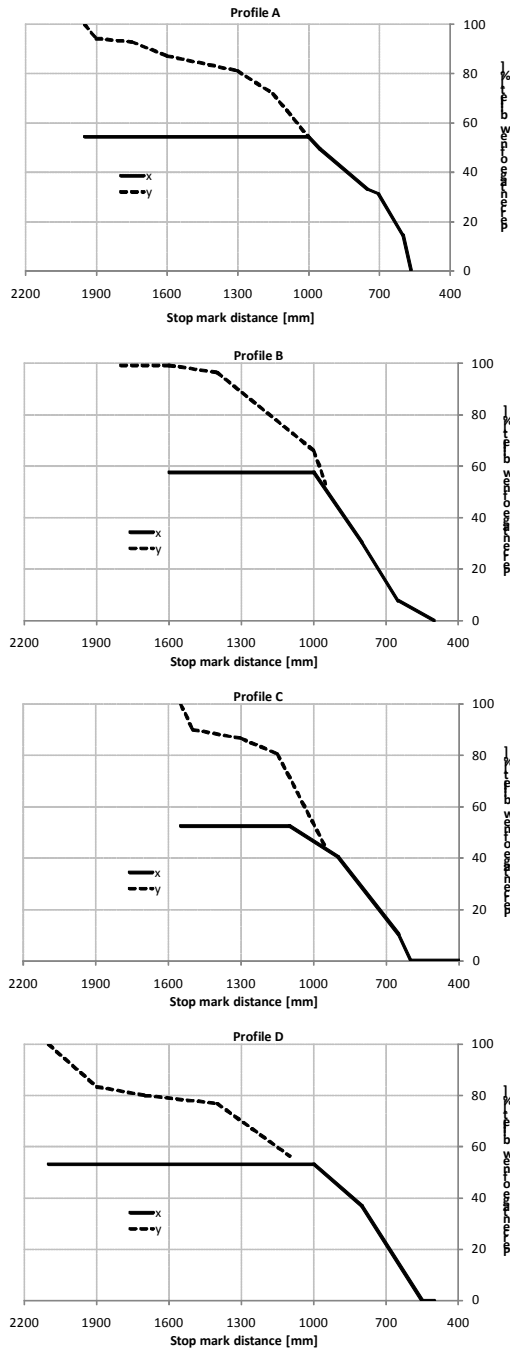


Figure 9. Percentage of new billet material, divided for X and Y flows, in the profiles over stop mark distance.

As can be observed, the replacement of the old billet with the new one (that is included within the dotted line in Fig. 8) firstly involved the thicker part of the profile (X), due to the combination of a more massive flow and proximity to the centre of the billet where the material flowed faster, and then the thinner part (Y). This is emphasized in Figure 9 where the percentage of the new billet over the stop mark distance for the X and Y parts of the profiles is shown. Graph representation allows an immediate overview of the charge weld over a single profile as well as between profiles. It is notable how for all the profiles the charge weld starts about at the same point (550mm) and always at the more massive part, which reach almost the whole filling with the new billet material before the charge weld in the Y flow is visible.

Table 3 reports the precise starting and ending measurements of the charge weld. The data are also reported for the each of the two flows.

As can be observed, only for profiles B and C (Fig. 10) there is a contemporary presence of the charge weld in the two flows, viceversa for profiles A and D, which show about 100 mm between the exhausting of the upper weld and the onset of the lower one. This is also in good agreement with the sensibly longer extent of weld for this two profiles, that are located in the upper part of the die. The location of the profiles on the die, together with their orientation, can therefore magnify the velocity unbalance effect, which has to be account for the charge weld extent and onset differences.

Table 3. Charge weld starting and ending points.

Profile	Start-stop [mm]	X [mm]	Y [mm]
A	1950-565	1000-565	1950-1090
B	1600-500	1000-500	1600-900
C	1550-400	1150-400	1550-950
D	2100-500	1000-400	2100-1080

3. NUMERICAL ANALYSES

FE model

The evolution of the transverse welds in the investigated extruded profiles was simulated by means of a purposely developed tool included in Altair HyperXrude® (Altair Engineering, Inc., Italy). It is a powerful commercial FE simulation code for the analysis and design of extrusion process and dies based on a fluid dynamic approach for modelling incompressible flows, including complex non-Newtonian fluid behaviour (Reddy et al. (1991)).

According to this type of modelling technique, the model of the aluminium billet had to copy the final shape of the extruded profile. It was therefore derived as the Boolean subtraction between a cylinder and the 3D CAD model of the die (Fig. 11a). The bearing lines were extracted by the 3D CAD model of the die and then projected on the billet surfaces (Fig. 11b).

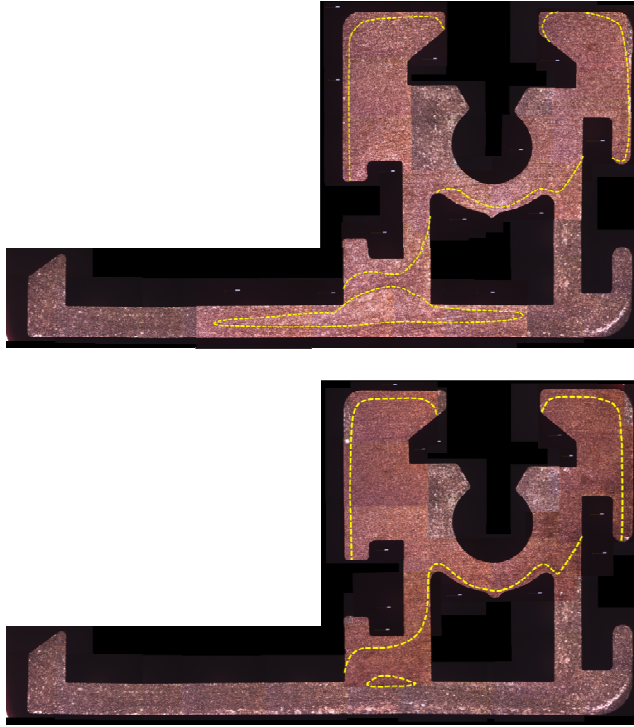


Figure 10. (top) Micrographical map of profile B at 1000mm and (bottom) of profile C at 950 mm from stop mark.

Due to the symmetry of the profile, only half of the model was simulated including profiles A and B. The resulting model was divided in 4 components to due the different boundary conditions involved: billet, porthole, that included the welding chamber, bearing and profile (Fig. 12). The computation of the transverse welds evolution did not require the simulation of the tools that were not included in the model. However, the heat exchange at the tool-billet interfaces was accounted for by setting proper values of the convective coefficients and reference temperatures for a third kind of boundary conditions for the heat diffusion equation at the surfaces.

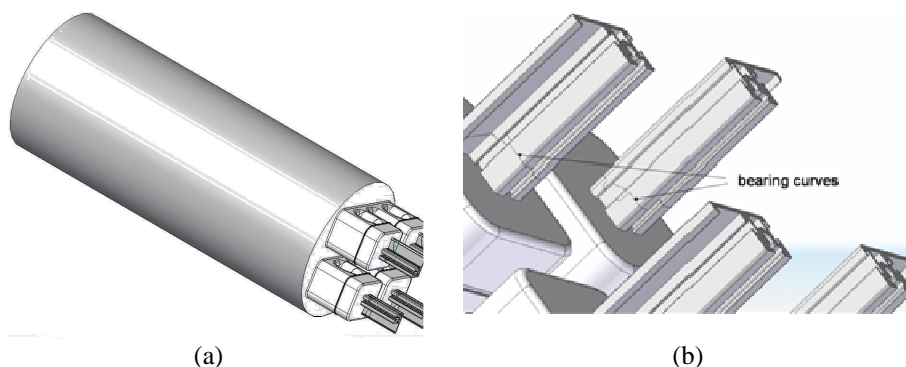


Figure 11. (a) 3D CAD model of the billet as required by HyperXrude®; (b) enlargement of the bearing curves definition.

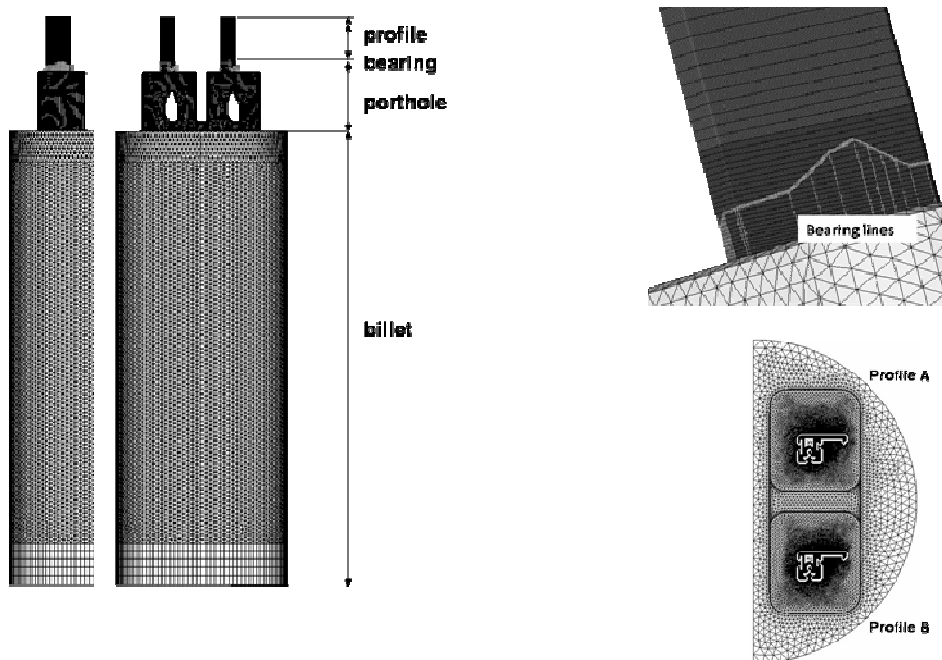


Figure 12. 3D FE model of the billet with the components definition.

The billet and porthole components were free meshed with 3D tetrahedral 4-noded elements. Due to the higher accuracy required in the bearing and profile regions, 3D prismatic 6-noded elements were used for these regions. The final model consisted of 1323583 elements and 538026 nodes.

The temperature of the billet at the exit of the preheating oven was set as the initial values in the FE simulation (i.e. 478°C). Among the four extruded billets, the first set of process parameters of the last billet recorded during the experimental campaign were used as the reference values for the ram speed setting and the results comparison (i.e. recording point 5.1 in Tab. 2).

At the four exits, zero normal stress boundary conditions were specified. This amounts to specifying zero pressure at the exits. A full sticking condition was set at the billet surfaces that, in the experimental set-up, were in contact with the container and die (porthole and welding chamber). For these, a convective coefficient of $3000 \text{ W/m}^2\text{C}^\circ$ was used; the references values for the heat equation were the respective initial temperatures of the tools (Tab. 1). Heat flux was set equal to zero for the free surface condition at the profile surfaces while in the bearing components a viscoplastic friction model with a coefficient value of 0.3 was used accounting for the material sliding that occur in these regions. The ram speed was imposed by applying the prescribed velocity of 6.2 mm/s at the billet back. The scheme of the applied boundary conditions is reported in Fig. 13.

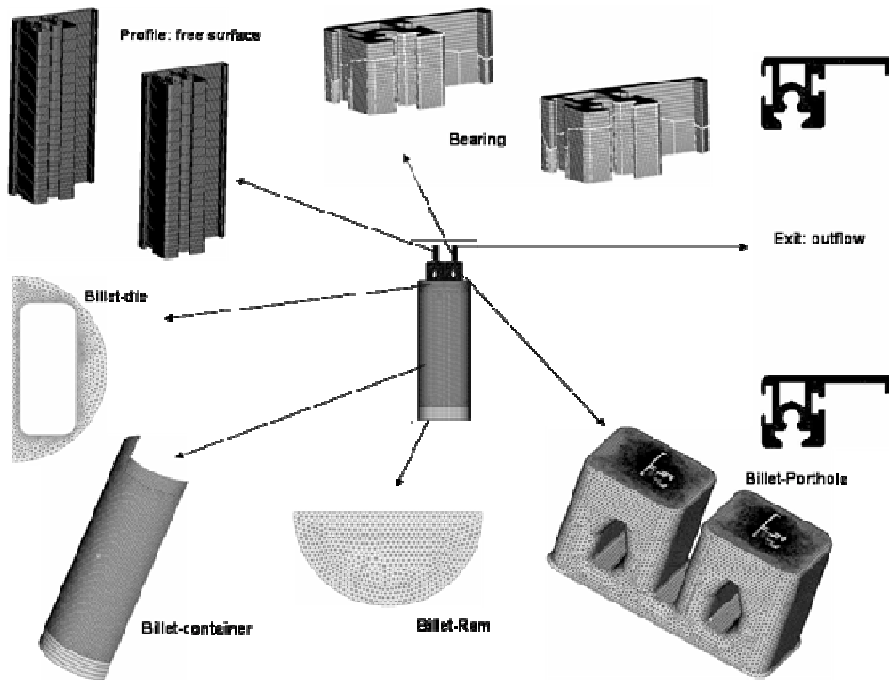


Figure 13. The boundary conditions applied to the 3D FE model of the extruded billet.

The flow of aluminum at elevated temperature exhibits a viscoplastic behaviour and is thus non-Newtonian. The effective flow stress of the AA6060 alloy was expressed by the Sellars-Tegart inverse sine hyperbolic model (Sellars et al. (1972)) then modified by Sheppard and Wright (Sheppard et al. (1979)) to yield the steady-state effective deviatoric flow stress:

$$\bar{\sigma} = \frac{1}{\alpha} \sinh^{-1} \left[\left[\left(\frac{Z}{A} \right) \right]^{\frac{1}{n}} \right] = \frac{1}{\alpha} \sinh^{-1} \left[\left[\frac{1}{A} \cdot \varepsilon \cdot \exp \left(\frac{Q}{RT} \right) \right]^{\frac{1}{n}} \right] \quad (1)$$

in which Z is the Zener-Hollomon parameter (Zener et al. (1944)) $n=3.515$, $Q=144000\text{J/mol}$, $A=5.91052 \cdot 10^9 \text{s}^{-1}$, $R=8.314 \text{ J/(K}^\circ\text{.mol)}$, $\alpha=3.464 \cdot 10^{-8} \text{ m}^2/\text{N}$ and T is expressed in K° .

The properties used for the AA6060 alloy are reported in Tab. 4:

Table 4. Physic, thermal and mechanical properties of the AA6060 used in the present study.

Density (Kg/m^3)	2685
Specific heat (J/Kg.K)	878
Conductivity (W/m.K)	200
Coefficient of thermal expansion ($1/\text{K}$)	1e-5
Young modulus (MPa)	40000
Poisson ratio	0.35

The scrap (weld length) calculation was performed by means of a transient analysis with moving boundaries. In this type of problem, the boundary conditions for the flow and heat transfer equations are treated as time-dependant and the position of the billet back and of the billet-container interface tracked during the simulation time. The mesh in the profile, bearing, porthole and welding chamber remain fixed, but in the billet region the elements scale down linearly in the extrusion direction at each time step. Altair HyperXrude® uses an Arbitrary Lagrange-Eulerian description of motion in problems with moving boundaries since it coalesces a good numerical stability and convergence properties with a robust mesh moving technique.

A variable decreasing number of time steps was defined starting from the first 10 seconds, in which the transverse welds were supposed to evolve and the velocity and thermal fields to reach the regime values, to the remaining 70 sec of simulation, used to evaluate the load-stroke history. The maximum number of nonlinear iteration was set equal to 25 with a convergence tolerances on velocities taken to be 10^{-3} . Even the tolerance for iterative equation solvers was set at 10^{-3} for velocities. A 90% of the deforming work was considered to be converted in heat.

Numerical results

The total simulation time was 132 hours. In Fig. 14 is shown the evolution of the transition between the new (grey) and the old (black) billets for the two profiles. As can be observed, the FE code correctly predicted the replacement of the billet that first involved the massive part (X) of the profile (steps 12 to 14) and then the thinner part (Y).

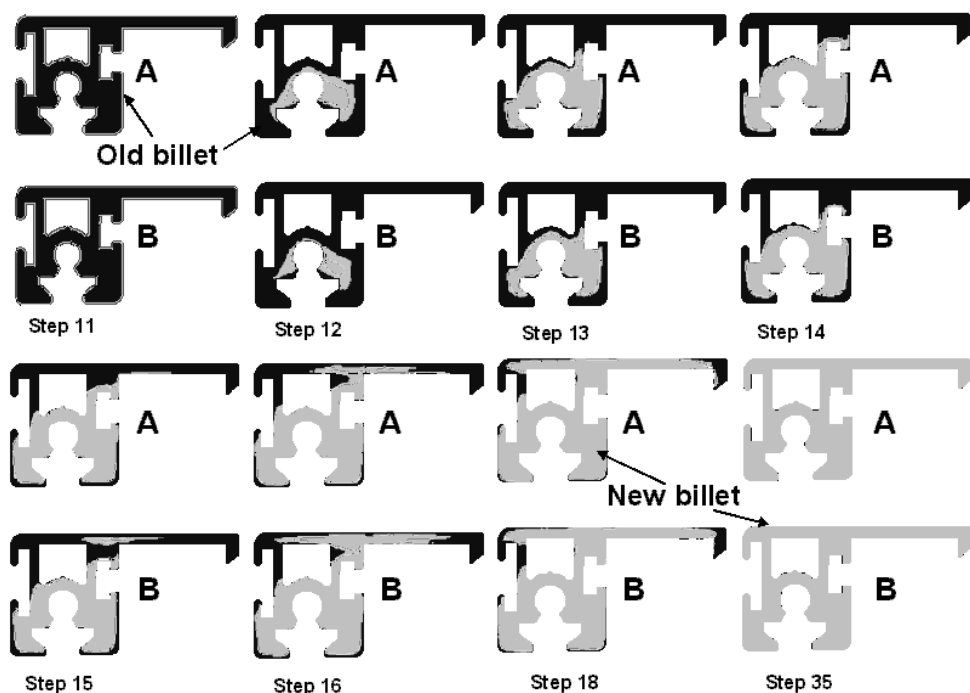


Figure 14. Evolution of the transverse weld at various subsequent time steps.

In addition, as experimentally observed, the code was found able to capture the faster filling of new material in the thicker part for the profile A than in B and the contrary for the thinner part.

In Figure 15 shows the percentage of the new billet as a function of the stop mark distance with the detail of the replacement of the new to the old billet for four simulation steps.

Minimal differences were observed between the prediction of the profiles A and B charge weld evolution, especially at the beginning of the replacement (Fig. 16).

4. DISCUSSION

A good agreement was found between the experimental results and numerical prediction in terms of profile exit velocity with a peak error at the steady state was less than 3% (Fig. 17a). A similarly accordance was observed for the profile exit temperature after the initial experimental thermal transient. Indeed, after the first 40 sec of extrusion, the peak difference was 7.8% (Fig. 17b).

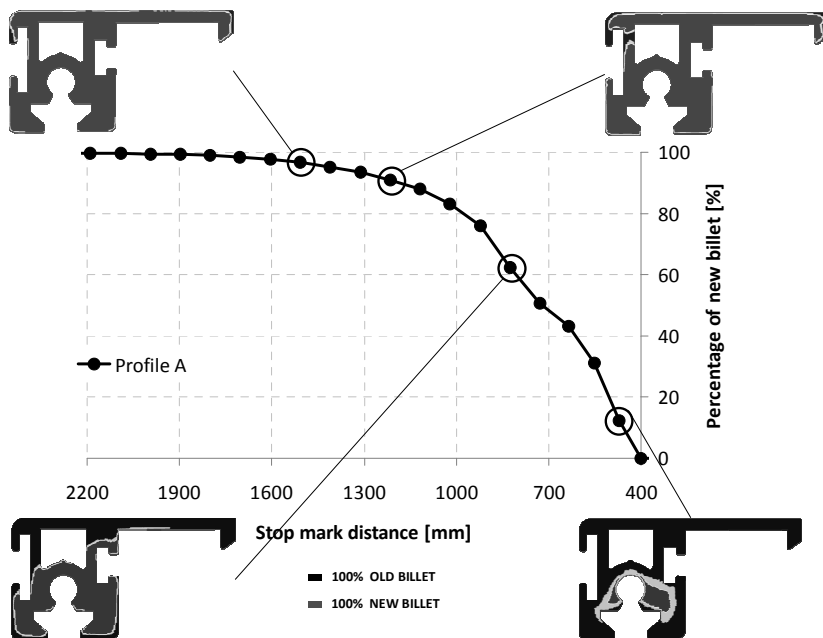


Figure 15. Simulated percentage of new material over stop mark distance

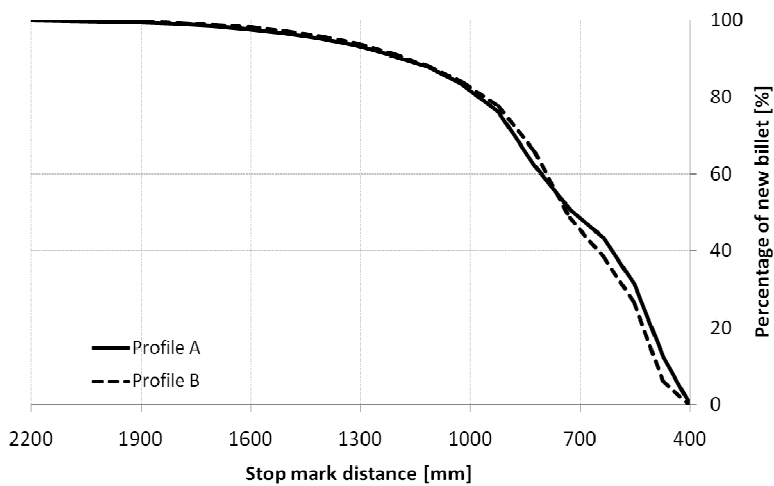


Figure 16. Comparison of the computed percentage of the new material history for the profiles A and B.

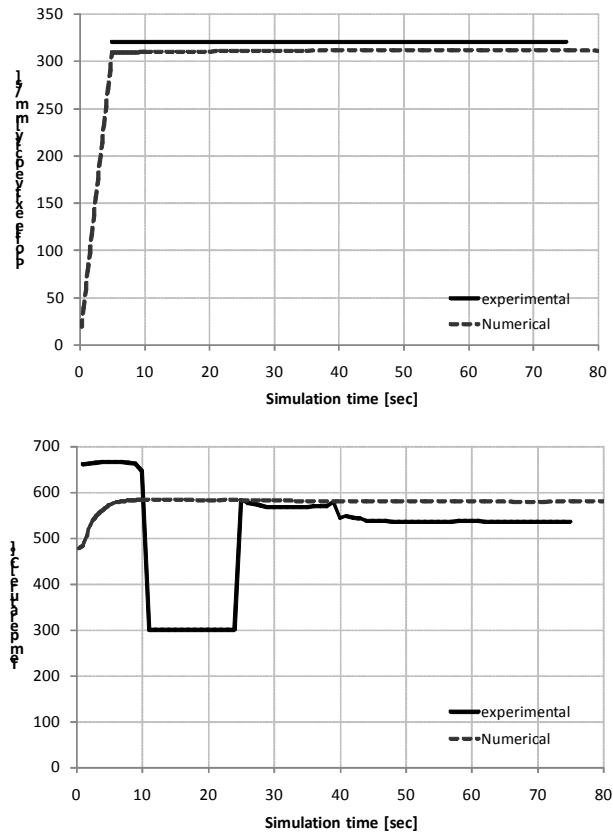


Figure 17. Comparison of the experimental vs. numerical results for the profile A in terms of exit velocity (left) and temperature (right).

A more detailed analysis over the charge weld evolution showed a very good agreement on the trend of profile filling by the new material. Figure 18 reports a comparison of numerical and experimental results for profiles A and B. The discrepancy of 200 mm between the experimental and numerical curves could be explained with the upsetting of the billet that is not computed by the FE code, or with the fact that the percentage of new material is computed at the final surface of the profile placed 50 mm after the die exit. However, more investigations are required in this direction.

In addition, Fig. 14 evidences how the numerical model was able to predict the faster filling of the material for massive part as experimentally observed (Tab. 2). It was also possible to highlight the overlapping of the weld for the two flows X and Y in profile B and the corresponding delay in profile A (step 15 of Figure 14). Even in this case the code predictions were in agreement with the experimental observation.

With the results showed above and a comparison with the industrial process parameters, it was possible to perform an evaluation of the die-process efficiency. First, the company was scrapping 2000mm by rule of thumb after the stop mark for weld contamination, which is a 8.7% of billet material. Above results showed that 2100mm is the extent that grant complete weld scrap, that is a 9.1%.

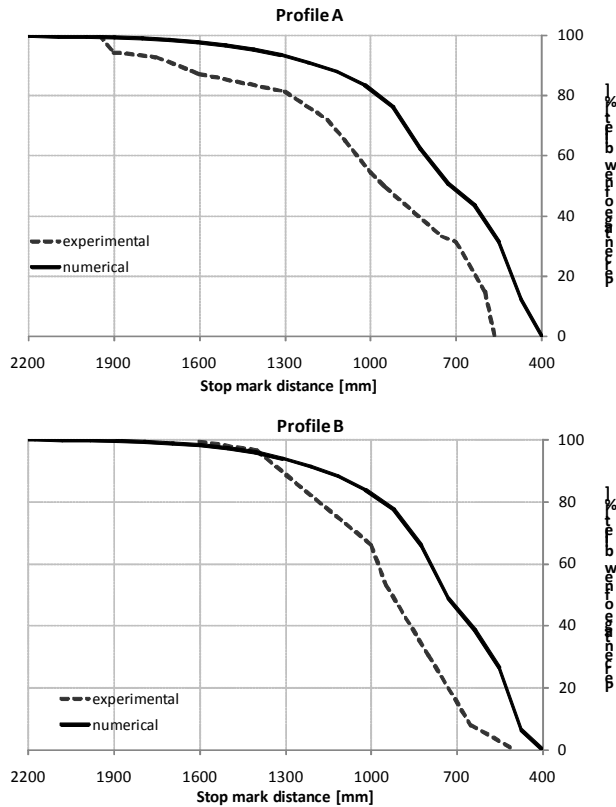


Figure 18. Comparison of the experimental vs. numerical results for percentage of new material over the stop mark distance for profile A (left) and B (right).

Moreover, it was normally scrapped 1000mm of profile before the charge weld for skin contamination while keeping a butt of 20mm; these adding up to 8.7% of billet material. Experimentally no trace of skin contamination was found in the profiles neither before than after the stop mark, thus suggesting that no scrap was necessary before the stop mark and that butt length could be reduced to 12mm, hence reducing the material waste to 2.6% of billet material. Adding up the above results, the company process scrap index of 17.4% could be lowered to the die-process efficiency index of 11.7% (retrieved by experimental activities).

5. CONCLUSIONS

Three typical extrusion problems have been investigated for a multi-hole die in a industrial environment. Most of the attention was directed to the charge weld as the seam weld would disappear within 250mm. The charge weld was investigated by means of microscopical analysis both in terms of extent, onset and evolution trend thus comparing the different die hole influence. Aside welds investigation, skin and butt defect were also inquired. In

addition, a numerical model was generated by means of the FE code HyperXrude® with the aim to evaluate the code capability to predict the charge weld evolution.

The comparison between experimental and numerical results showed a good agreement both in terms of general trend and exhausting points of the percentage of new material over the stop mark distance. This encourage the extension of the work in order to include the back end as the seam weld computation within the FE model.

Finally, a synthetic index was defined to give a measure of the die and process efficiency in terms of reduced profile scraps. This can be used to compare two different die designs and theoretical and practical process parameters (e.g. butt and scraps length).

Acknowledgments

The authors would like to thank Eng. Floriano Bagagli of Alutitan s.p.a. San Marino, Italy for the interest, support and for his help to perform the experiments.

REFERENCES

- [1] N.C. Parson, J.D. Hankin, A.J. Bryant, 1992. "The metallurgical background to Problems Occurring During the Extrusion of 6XXX alloys" *Proceedings of 5th International Aluminum Extrusion Technology Seminar*, Chicago, USA, vol II, pp.13-24;
- [2] M. Lefsta, O. Reiso, V. Johnsen, 1992. "Flow of the billet Surface in Aluminum Extrusion" *Proceedings of 5th International Aluminum Extrusion Technology Seminar*, Chicago, USA, vol II, pp.503-518;
- [3] H. Valberg 1988, "Physical Simulation of Metal Extrusion by Means of Model Materials" *Proceedings of 4th International Aluminum Extrusion Technology Seminar*, Chicago, USA, vol II, pp.321-328;
- [4] H. Valberg, 2010. "Understanding Metal Flow In Aluminium Extrusion By Means Of Emptying Diagrams", *International Journal of Material Forming*, Vol. 3, issue 1, pp. 391-394;
- [5] B. Reggiani, L. Donati, L. Tomesani, 2010. Effect of different FE simulation codes in the stress analysis of extrusion dies. In the *Proc. of the 13th International ESAFORM Conference on Material Forming*, University of Brescia, Brescia, Italy, 7-9 April 2010.
- [6] L. Tomesani, B. Reggiani, L. Donati, N. B. Khalifa, A. E. Tekkaya, 2010. Experimental evaluation and numerical analysis of speed-related effects in aluminium extrusion. In the *Proceedings of the 60th CIRP, Accademia Internazionale di Ingegneria della Produzione* (<http://www.cirp.net>), 22nd-28th August, Pisa, Italy.
- [7] J. Zhou, 2003. "3D FEM simulation of the whole cycle of aluminium extrusion throughout the transient state and the steady state using the updated Lagrangian approach", *J. Mater. Proc. Technol.*, 134, pp. 383-397.
- [8] C. Sommitsch, R. Sievert, T. Wlanis, B. Günther and V. Wieser, 2007 "Modelling of creep-fatigue in containers during aluminium and copper extrusion", *Computational Materials Science*, 39, pp. 55-64.
- [9] B. Bourqui, A. Huber, C. Moulin, A. Bunetti, 2002. "Improved weld seam quality using 3D FEM simulations in correlation with practice", in: *Proceedings of the First EAA, European Aluminum Association* Vol. 27-2 - Ed. Metallurgical Science and Technology 29 Extruders Division, Brescia, Italy.

- [10] L. Donati, L. Tomesani, 2004. "The prediction of seam welds quality in aluminum extrusion". *Journal of Materials Processing Technology*, 153–154, pp. 366–373.
- [11] L. Donati, L. Tomesani, 2005. "The effect of die design on the production and seam weld quality of extruded aluminum profiles". *Journal of Materials Processing Technology*, 164–165, pp.1025–1031;
- [12] L. Donati, L. Tomesani, G. Minak, 2007. "Characterization of seam weld quality in AA6082 extruded profiles". *Journal of Materials Processing Technology*, 191, pp. 127–131.
- [13] G. Liu, J. Zhou, J. Duszczek, 2008. "FE analysis of metal flow and weld seam formation in a porthole die during the extrusion of a magnesium alloy into a square tube and the effect of ram speed on weld strength", *J. Mater. Process. Technol.*, 200,185 – 198.
- [14] Reddy, M.P., Reddy, J.N., Akay, H.U., 1991. "A 3-D penalty finite element model of forming processes. Advances in Finite Element Analysis in Fluid Dynamics", *ASME*, 123, pp. 61-76
- [15] Sellars, C.M., Tegart, W.J.McG., 1972. "Hot workability", *International Metallurgical Reviews*, 17(1), pp. 1-24
- [16] Sheppard, T., Wright, D., 1979. "Deformation of flow stress: Part 1 constitutive equation for aluminium alloys at elevated temperature". *Metals Technology*, 6, pp. 215-223.
- [17] Zener, C., Hollomon, J.H., 1994. "Effect of strain-rate upon plastic flow of steel". *Journal of Applied Physics*, 15, pp. 22-32

AGGIORNAMENTO AUTOMATICO DI UN MODELLO AD ELEMENTI FINITI B-SPLINE MEDIANTE RILIEVI SPERIMENTALI DI FUNZIONI DI RISPOSTA IN FREQUENZA

Antonio Carminelli

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia (DIEM),

Università di Bologna, Italia

E-mail: antonio.carminelli@unibo.it

Giuseppe Catania

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia (DIEM),

Università di Bologna, Italia

E-mail: giuseppe.catania@unibo.it

Sommario. Viene presentata una procedura di aggiornamento di un modello ad Elementi Finiti (EF) di componenti in parete sottile basata su funzioni di forma B-spline, utilizzando Funzioni di Risposta in Frequenza (FRF) come dati di ingresso. La procedura si basa sulla minimizzazione di una funzione obiettivo definita mediante la somma dei quadrati della differenza tra la risposta del modello e quella misurata sperimentalmente alla stessa frequenza. I parametri ottimizzati sono i coefficienti dei modelli di vincolo e di smorzamento, entrambi descritti facendo uso di funzioni B-spline. L'incompatibilità tra i gdl misurati e quelli del modello EF è stata risolta mediante una trasformazione di coordinate che fa uso delle stesse funzioni di forma del modello EF. La correttezza della procedura è stata testata tramite un esempio numerico relativo ad una piastra quadrata vincolata su di un lato. Le FRF utilizzate come dati di ingresso della procedura sono state generate numericamente. I risultati sono presentati e discussi.

Keywords: *updating, B-spline, elementi finiti, FRF, vincoli, smorzamento.*

1. INTRODUZIONE

I modelli ad EF possono essere utilizzati per simulare la risposta di una struttura o di un componente meccanico in modo accurato. Recentemente sono stati pubblicati diversi lavori sull'utilizzo di funzioni di tipo spline come funzioni di forma per modelli EF di tipo tridimensionale [1], bidimensionale [2] e monodimensionale [3]. In [4] è stato mostrato come tali modelli ad EF possano mostrare una maggiore accuratezza, rispetto ai classici modelli ad EF impieganti funzioni polinomiali, soprattutto per ciò che riguarda la simulazione della risposta vibratoria. Tuttavia, avendo a che fare con modelli teorici, è necessario considerare le incertezze del modello quali, ad esempio, quelle legate alla modellazione di vincoli e giunti, alle proprietà del materiale o al modello di smorzamento.

Nonostante la letteratura riguardante le tecniche di aggiornamento (o updating) [5-10] di modelli ad EF sia vasta, alcune tematiche sono ancora irrisolte. L'utilizzo di parametri modali, di FRF oppure di altre tipologie di risposte come dati d'ingresso della procedura di updating è un argomento dibattuto. La tecnica di parametrizzazione è una problematica rilevante che deve essere trattata per evitare il mal-condizionamento delle equazioni: è importante che i parametri di updating siano in grado di modellare la risposta del sistema evitando al contempo soluzioni non fisiche. Inoltre l'aggiornamento dei modelli delle condizioni di vincolo [6] e dello smorzamento strutturale [9], a conoscenza degli scriventi, sono argomenti non estesamente trattati in letteratura. Infine, l'incompatibilità tra i gradi di libertà (gdl) di misura e i gdl del modello EF è una problematica ricorrente, e tecniche di espansione o riduzione sono generalmente utilizzate per risolvere tale tipologia di problema [5]. Un approccio più generale, tuttavia, dovrebbe prendere in considerazione anche l'utilizzo di diverse tipologie di gdl nei due modelli [12].

In questo lavoro vengono utilizzate FRF come dati di ingresso di una procedura di updating di un modello ad EF per componenti in parete sottile basata su funzioni di forma B-spline. La procedura si basa sulla minimizzazione di una funzione obiettivo definita dalla somma dei quadrati dei residui, definiti come la differenza tra la risposta del modello e quella misurata sperimentalmente alla stessa frequenza. I parametri ottimizzati sono i coefficienti dei modelli di vincolo e di smorzamento, entrambi descritti facendo uso di funzioni B-spline. Il modello di smorzamento è definito su un intervallo di frequenza più esteso della banda in cui sono disponibili le misure, al fine di considerare anche l'effetto di eventuali modi prossimi all'intervallo di frequenza misurato. L'incompatibilità tra i gdl misurati e quelli del modello ad EF è stata risolta mediante una trasformazione di coordinate che fa uso delle stesse funzioni di forma del modello ad EF.

La procedura è stata testata utilizzando il modello di una piastra quadrata con i modelli di vincolo e smorzamento proposti. Le FRF sono state generate numericamente utilizzando lo stesso modello di piastra ma con differenti modelli di vincolo e smorzamento. I risultati vengono presentati e discussi.

2. ELEMENTO FINITO GUSCIO B-SPLINE

Modello guscio B-spline

La geometria di un componente in parete sottile con superficie media a doppia curvatura può essere efficacemente descritta utilizzando una mappa dallo spazio parametrico ξ, η, τ $0 \leq \xi, \eta, \tau \leq 1$ allo spazio Euclideo tridimensionale (x,y,z). Il vettore posizione della superficie media di un componente in parete sottile, in un sistema di riferimento cartesiano fisso $\mathbf{O}\{x,y,z\}$, può essere definito per mezzo del prodotto tensoriale di funzioni B-spline [13]:

$$\mathbf{r}(\xi, \eta) = \begin{Bmatrix} r_x \\ r_y \\ r_z \end{Bmatrix} = \sum_{i=1}^m \sum_{j=1}^n B_i^p(\xi) \cdot B_j^q(\eta) \cdot \mathbf{P}_{ij} \quad (1)$$

dove:

- \mathbf{P}_{ij} sono $m \times n$ punti di controllo (pdc) che descrivono la superficie media B-spline;
- $B_i^p \xi$ (e rispettivamente $B_j^q \eta$) è la i -esima (j -esima) funzione B-spline normalizzata di grado p (q) definita rispetto alla coordinata curvilinea $\xi(\eta)$ per mezzo del vettore dei nodi $\mathbf{U} = \xi_1, \dots, \xi_r$ ($\mathbf{V} = \eta_1, \dots, \eta_r$).

L'elemento finito noto come *guscio degenero* è utilizzato in molti software commerciali per la sua semplicità. Il vettore posizione dell'elemento guscio degenero può essere espresso come:

$$\mathbf{s}(\xi, \eta, \tau) = \sum_{i=1}^m \sum_{j=1}^n B_i^p(\xi) \cdot B_j^q(\eta) \cdot \left[\mathbf{P}_{ij} + t_{ij} \left(\tau - \frac{1}{2} \right) \mathbf{v}_{ij}^3 \right] \quad (2)$$

dove i vettori \mathbf{v}_{ij}^3 e i coefficienti t_{ij} possono essere calcolati seguendo la tecnica di interpolazione definita in [14].

La funzione di spostamento può essere definita seguendo l'approccio isoparametrico [15]:

$$\begin{aligned} \mathbf{d}(\xi, \eta, \tau) &= \sum_{i=1}^m \sum_{j=1}^n B_i^p(\xi) \cdot B_j^q(\eta) \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} t_{ij} \left(\tau - \frac{1}{2} \right) \begin{bmatrix} -\mathbf{v}_{ij}^2 & \mathbf{v}_{ij}^1 \end{bmatrix} \cdot \begin{Bmatrix} u_{ij} \\ v_{ij} \\ w_{ij} \\ \alpha_{ij} \\ \beta_{ij} \end{Bmatrix} = \\ &= \begin{bmatrix} \mathbf{N}_u \\ \mathbf{N}_v \\ \mathbf{N}_w \end{bmatrix} \cdot \boldsymbol{\delta} = \mathbf{N} \cdot \boldsymbol{\delta}, \end{aligned} \quad (3)$$

dove $\boldsymbol{\delta}$ è il vettore contenente $N = 5 \cdot m \cdot n$ gdl generalizzati:

$$\boldsymbol{\delta}^T = u_{11} \quad v_{11} \quad w_{11} \quad \alpha_{11} \quad \beta_{11} \quad \cdots \quad u_{mn} \quad v_{mn} \quad w_{mn} \quad \alpha_{mn} \quad \beta_{mn}, \quad (4)$$

u_{ij} , v_{ij} e w_{ij} sono gdl associati alla modellazione della traslazione, α_{ij} e β_{ij} sono gdl associati alla modellazione della rotazione, $\mathbf{v}_{ij}^1, \mathbf{v}_{ij}^2, \mathbf{v}_{ij}^3$ è una terna ortonormale definita su \mathbf{P}_{ij} a partire dal vettore \mathbf{v}_{ij}^3 [14].

Le equazioni del moto possono essere ottenute utilizzando il principio di minimo dell'energia potenziale totale [15]:

$$\Pi = U + W \rightarrow \min \quad (5)$$

dove U è il potenziale interno del sistema e W è il potenziale delle forze esterne,

includendo quello delle forze d'inerzia. Seguendo la trattazione esposta in [12] si ottengono le equazioni del moto:

$$\mathbf{M} \cdot \ddot{\boldsymbol{\delta}} + \mathbf{K}_f \cdot \boldsymbol{\delta} = \mathbf{F} \quad (6)$$

dove \mathbf{M} è la matrice di massa, \mathbf{K}_f è la matrice di rigidezza del sistema non vincolato e \mathbf{F} è il vettore delle forze esterne applicate.

Modello dei vincoli

I vincoli sono stati imposti considerando nel funzionale dell'energia potenziale totale Π anche l'energia potenziale ΔW delle forze vincolari per unità di superficie $\boldsymbol{\Phi}_c$, considerate applicate sulla superficie esterna dell'elemento guscio:

$$\Delta W = -\frac{1}{2} \int_S (\mathbf{d}^T \cdot \boldsymbol{\Phi}_c) dS = \frac{1}{2} \boldsymbol{\delta}^T \cdot \int_S \mathbf{N}^T \cdot \mathbf{R} \cdot \mathbf{N} dS \cdot \boldsymbol{\delta} \quad (7)$$

con:

$$\boldsymbol{\Phi}_c = -\mathbf{R} \cdot \mathbf{d}, \quad (8)$$

dove \mathbf{R} è una matrice 3x3 che contiene i coefficienti di rigidezza r_{ab} ($a, b=1,2,3$) del modello di vincolo elastico distribuito, descritto utilizzando delle funzioni B-spline:

$$r_{ab} = \sum_{i=1}^{m^{ab}} \sum_{j=1}^{n^{ab}} B_i^{p^{ab}} \cdot B_j^{q^{ab}} \cdot \kappa_{ij}^{ab} \quad (9)$$

dove $B_i^{p^{ab}}$ e $B_j^{q^{ab}}$ sono funzioni B-spline monovariate normalizzate, definite sui vettori dei nodi \mathbf{U}^{ab} e \mathbf{V}^{ab} .

Includendo l'energia potenziale ΔW delle forze vincolari (Eqn. (7)) nell'espressione dell'energia potenziale totale Π e imponendone la stazionarietà [12], le equazioni del moto diventano:

$$\mathbf{M} \cdot \ddot{\boldsymbol{\delta}} + \mathbf{K}_f + \Delta \mathbf{K} \cdot \boldsymbol{\delta} = \mathbf{F}, \quad (10)$$

dove la matrice di rigidezza delle forze vincolari è:

$$\Delta \mathbf{K} = \int_S \mathbf{N}^T \cdot \mathbf{R} \cdot \mathbf{N} dS. \quad (11)$$

Modello dello smorzamento

Per strutture poco smorzate può essere lecito utilizzare l'ipotesi di smorzamento reale (autovettori reali). L'ipotesi di smorzamento reale è stata imposta considerando un termine viscoso nelle equazioni del moto:

$$\mathbf{M} \cdot \ddot{\boldsymbol{\delta}} + \mathbf{C} \cdot \dot{\boldsymbol{\delta}} + \mathbf{K}_f + \Delta \mathbf{K} \cdot \boldsymbol{\delta} = \mathbf{F} \quad (12)$$

dove la matrice di smorzamento \mathbf{C} è:

$$\mathbf{C} = \mathbf{\Phi}^{-T} \cdot \mathbf{C}_d \cdot \mathbf{\Phi}^{-1} \quad (13)$$

e

$$\mathbf{C}_d = \begin{bmatrix} 2\zeta_1\omega_1 & 0 & \dots & 0 \\ 0 & 2\zeta_2\omega_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 2\zeta_N\omega_N \end{bmatrix}, \quad (14)$$

dove $\mathbf{\Phi}$ è la matrice degli autovettori $\mathbf{\Phi}_i$ ottenuti risolvendo l'autoproblema generalizzato

$$\mathbf{K} - \omega_i^2 \mathbf{M} \mathbf{\Phi}_i = \mathbf{0}, \quad (15)$$

ω_i^2 è l'i-esimo autovalore dell'Eqn. (15). ζ_i è l'i-esimo fattore di smorzamento modale che può essere valutato tramite la seguente formula:

$$\zeta_i = \zeta f_i = \zeta 2\pi \cdot \omega_i, \quad (16)$$

dove lo smorzamento in funzione della frequenza ζf è descritto come combinazione lineare dei coefficienti di controllo γ_z e delle funzioni B-spline B_z definite su un vettore di nodi uniformemente spaziatto:

$$\zeta f = \zeta f(u) = \sum_{z=1}^{n_z} B_z(u) \cdot \gamma_z; \quad f = f_{ST} + u \cdot f_{FI} - f_{ST}; \quad u \in [0,1], \quad (17)$$

dove f_{ST} e f_{FI} sono, rispettivamente, il limite inferiore e superiore dell'intervallo di frequenze sul quale il modello di smorzamento è definito. Al di fuori di tale intervallo di frequenze, per il calcolo della matrice \mathbf{C} del modello ad elementi finiti, è stato considerato un valore costante per i fattori di smorzamento. Inoltre, per poter considerare l'effetto dei modi al di fuori dell'intervallo di frequenza misurato, è stato imposto $f_{FI} = c \cdot f_{MAX}$ (con $c > 1$), dove con f_{MAX} si è indicato il valore della frequenza massima presente nei dati acquisiti.

Indicando con $\tilde{\cdot} = \mathcal{F}(\cdot)$ l'operatore di trasformazione secondo Fourier, l'equazione di equilibrio dinamico nel dominio della frequenza è:

$$-\omega^2 \mathbf{M} + j\omega \mathbf{C} + \mathbf{K}_r + \Delta \mathbf{K} \cdot \tilde{\mathbf{\delta}} = \mathbf{Z} \omega \cdot \tilde{\mathbf{\delta}} = \mathbf{H}^{-1} \omega \cdot \tilde{\mathbf{\delta}} = \tilde{\mathbf{F}}, \quad (18)$$

dove $\mathbf{Z} \omega$ è la matrice di impedenza dinamica e $\mathbf{H} \omega = \mathbf{Z} \omega^{-1}$ è la matrice delle ricettanze.

3. TECNICA DI UPDATING

La procedura di updating utilizza direttamente FRF come dati di ingresso e si basa sulla minimizzazione ai minimi quadrati dei residui, definiti come la differenza tra la FRF del modello e quella misurata sperimentalmente considerate alla stessa frequenza. I parametri

ottimizzati simultaneamente sono i coefficienti dei modelli di vincolo e di smorzamento introdotti nelle sezioni precedenti.

Le ℓ FRF $H_b^x \omega$, con $b=1, \dots, \ell$, utilizzate come dati di ingresso, sono state raccolte nel vettore $\mathbf{h}_x \omega$:

$$\mathbf{h}_x \omega = \begin{Bmatrix} H_1^x \omega \\ \vdots \\ H_\ell^x \omega \end{Bmatrix}. \quad (19)$$

Nessun elemento del vettore $\mathbf{h}_x \omega$ può essere direttamente confrontato con uno degli elementi della matrice $\mathbf{H} \omega$, perché il vettore $\tilde{\mathbf{\delta}}$ contiene gdl non fisici. Le FRF corrispondenti ai gdl fisici del modello possono essere ricavate utilizzando le funzioni di forma del modello ad EF: se si considera la forza applicata nel punto $\mathbf{P}^i = \mathbf{s}(\xi_i, \eta_i, \tau_i)$ lungo la direzione $\boldsymbol{\varphi}$ e la risposta misurata nel punto $\mathbf{P}^r = \mathbf{s}(\xi_r, \eta_r, \tau_r)$ lungo la direzione $\boldsymbol{\psi}$, la FRF analitica corrispondente è:

$$H_{\boldsymbol{\psi}, \boldsymbol{\varphi}}^{r,i} \omega = \mathbf{N}_{\boldsymbol{\psi}}(\xi_r, \eta_r, \tau_r) \cdot \mathbf{H} \omega \cdot \mathbf{N}_{\boldsymbol{\varphi}}^T(\xi_i, \eta_i, \tau_i), \quad (20)$$

dove $\boldsymbol{\varphi}$ e $\boldsymbol{\psi}$ possono assumere un valore tra \mathbf{u} , \mathbf{v} or \mathbf{w} (Eqn. (3)).

Associando ad ogni FRF presente nel vettore $\mathbf{h}_x \omega$ la corrispondenza $i, r, \boldsymbol{\varphi}, \boldsymbol{\psi}$, è possibile ricavare le relative FRF analitiche che possono essere raccolte ordinatamente nel vettore $\mathbf{h}_a \omega, \mathbf{p}$:

$$\mathbf{h}_a \omega, \mathbf{p} = \begin{Bmatrix} H_{\boldsymbol{\psi}, \boldsymbol{\varphi}}^{r,i} \omega, \mathbf{p} \\ \vdots \\ H_{\mathbf{0}, \boldsymbol{\sigma}}^{\ell, t} \omega, \mathbf{p} \end{Bmatrix}, \quad (21)$$

dove $\mathbf{p} = p_1 \cdots p_{n^p}^T$ è il vettore che contiene i parametri p_k da ottimizzare. Si noti che nessun limite è stato imposto sul numero di punti di misura o di applicazione della forzante.

In generale le FRF analitiche presenti nel vettore $\mathbf{h}_a \omega, \mathbf{p}$ sono funzioni che dipendono non-linearmente dai parametri p_k . Tale dipendenza è stata linearizzata, fissata la pulsazione ω_i , mediante espansione in serie di Taylor troncata nell'intorno di $\mathbf{p}=\mathbf{p}_0$:

$$\mathbf{h}_a \omega_i, \mathbf{p}_0 + \sum_{k=1}^{n^p} \frac{\partial \mathbf{h}_a \omega_i, \mathbf{p}_0}{\partial p_k} \Delta p_k \simeq \mathbf{h}_x \omega_i, \quad (22)$$

dove la sensibilità delle FRF $H_{\boldsymbol{\psi}, \boldsymbol{\varphi}}^{r,i}$ rispetto al generico parametro p_k è:

$$\begin{aligned}
\frac{\partial H_{\Psi, \Phi}^{r,i}}{\partial p_k} \omega, \mathbf{p} &= \mathbf{N}_{\Psi}(\xi_r, \eta_r, \gamma_r) \cdot \frac{\partial \mathbf{H}}{\partial p_k} \omega, \mathbf{p} \cdot \mathbf{N}_{\Phi}^T(\xi_i, \eta_i, \gamma_i) = \\
&= -\mathbf{N}_{\Psi}(\xi_r, \eta_r, \gamma_r) \cdot \mathbf{H} \omega, \mathbf{p} \cdot \frac{\partial \mathbf{Z}}{\partial p_k} \omega, \mathbf{p} \cdot \mathbf{H} \omega, \mathbf{p} \cdot \mathbf{N}_{\Phi}^T(\xi_i, \eta_i, \gamma_i)
\end{aligned} \quad (23)$$

L'Eqn.(22) può essere espressa come:

$$\left[\frac{\partial \mathbf{h}_a}{\partial p_1} \omega_i, \mathbf{p}_o, \dots, \frac{\partial \mathbf{h}_a}{\partial p_k} \omega_i, \mathbf{p}_o, \dots, \frac{\partial \mathbf{h}_a}{\partial p_{np}} \omega_i, \mathbf{p}_o \right] \begin{bmatrix} \Delta p_1 \\ \vdots \\ \Delta p_k \\ \vdots \\ \Delta p_{np} \end{bmatrix} = \mathbf{h}_x \omega_i - \mathbf{h}_a \omega_i, \mathbf{p}_o \quad (24)$$

o in maniera compatta:

$$\mathbf{S}_i \cdot \Delta \mathbf{p} = \Delta \mathbf{h}_i, \quad (25)$$

dove \mathbf{S}_i è la matrice delle sensibilità calcolata in corrispondenza della pulsazione ω_i e $\Delta \mathbf{p}$ è il vettore delle variazioni dei parametri.

Utilizzando i dati delle FRF per n^f valori della pulsazione ω_i , è possibile ottenere una stima ai minimi quadrati delle n^p variazioni Δp_k definendo la funzione errore \mathbf{e} :

$$\mathbf{e} = \sum_{i=1}^{n^f} \mathbf{S}_i \cdot \Delta \mathbf{p} - \Delta \mathbf{h}_i, \quad n^f \gg n^p, \quad (26)$$

e minimizzando la funzione quadratica g :

$$g = \mathbf{e}^T \cdot \mathbf{e} \rightarrow \min. \quad (27)$$

Il sistema di equazioni che ne deriva può essere mal-condizionato poichè le FRF possono avere sensibilità molto diverse a seconda del parametro p_k che si considera. Per limitare tale problema è stata utilizzata una normalizzazione delle variabili di updating. È stata inoltre utilizzata una trasformazione delle variabili per vincolare i valori dei parametri ottimizzati in intervalli compatti predefiniti senza l'utilizzo di ulteriori parametri [12].

La tecnica è iterativa e prevede ad ogni passo l'approssimazione del primo ordine effettuata nell'Eqn.(22).

4. SIMULAZIONI NUMERICHE

Viene presentato un esempio numerico relativo al modello di una piastra quadrata incastrata su un lato. Sia la geometria che il campo di spostamento sono descritti con funzioni B-spline di terzo grado e una griglia di 15x11 punti di controllo.

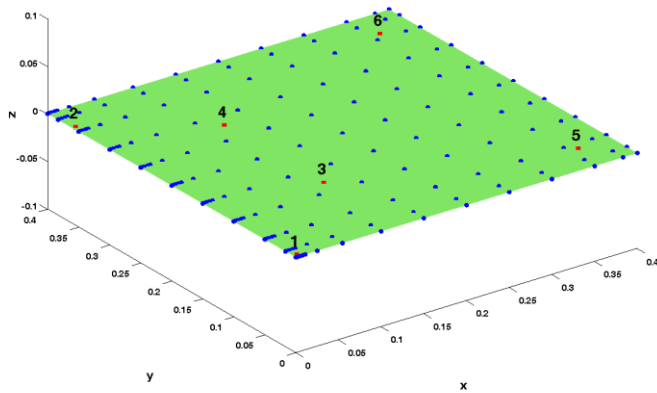


Figura 1. Modello ad EF della piastra quadrata con 15x11 pdc (in blu) e 6 gdl di misura (in rosso). I vincoli sono applicati sul lato $x=0$.

Il modello ad EF della piastra quadrata (lato $L=0.4$ m) è rappresentato in Fig.1. Il materiale è omogeneo ed isotropo con le seguenti proprietà:

$$E = 2.1 \cdot 10^{11} \text{ Pa}, \quad \nu_{12} = 0.3, \quad \rho = 7850 \text{ Kg} / \text{m}^3 \quad (28)$$

Simulazione delle FRF misurate

Per testare la correttezza della procedura di updating sono state utilizzate FRF generate numericamente con lo stesso modello di guscio, ma utilizzando una diversa descrizione per il modello di vincolo e per quello di smorzamento. I vincoli sono stati imposti su di un lato della piastra quadrata ($x=0$). I gdl di traslazione secondo gli assi x e y sono stati eliminati sul lato vincolato. I restanti gdl (traslazioni secondo l'asse z e rotazioni) sono stati vincolati utilizzando delle rigidità distribuite, agenti sul lato inferiore della piastra ($\tau = 0$), descritte mediante il modello proposto nel presente lavoro, dove gli elementi della matrice \mathbf{R} sono tutti nulli tranne r_{33} , per il quale valgono (Eqn.(9)):

- $m^{33}=7$
- $n^{33}=1$
- $p^{33}=2$
- $q^{33}=0$
- $\kappa_{ij}^{33} = 10^9 \cdot \begin{bmatrix} 10^3 & 1 & 1 & 10^3 & 1 & 1 & 10^3 \end{bmatrix}$

I fattori di smorzamento modali ζ_i utilizzati per generare la matrice \mathbf{C} sono riportati in Fig.2. Con tali valori sono state generate le FRF, nell'intervallo di frequenza $[0, 800]$ Hz, assumendo che la forzante sia applicata in corrispondenza del gdl 1 (quadrato rosso "1" in Fig.1) lungo la direzione z e simulando la risposta in corrispondenza di sei gdl (spostamento w in corrispondenza dei sei quadrati rossi in Fig.1).

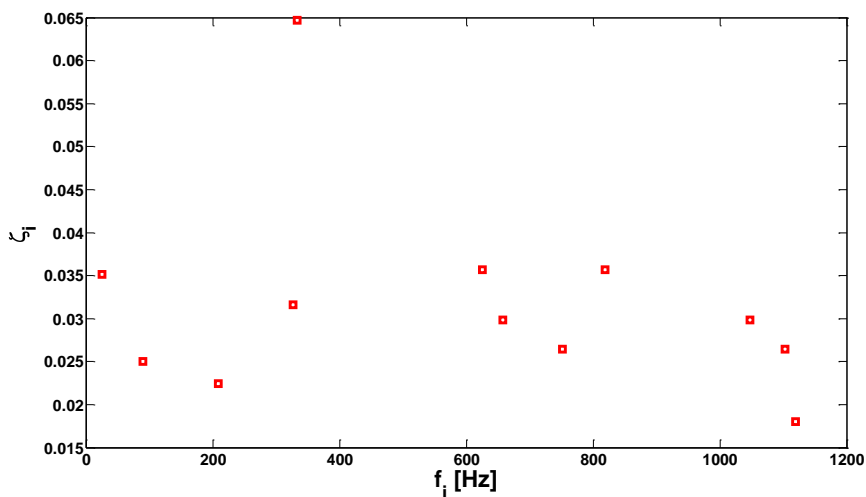


Figura 2. I fattori di smorzamento modali ζ_i utilizzati per generare le FRF utilizzate come input della procedura di updating.

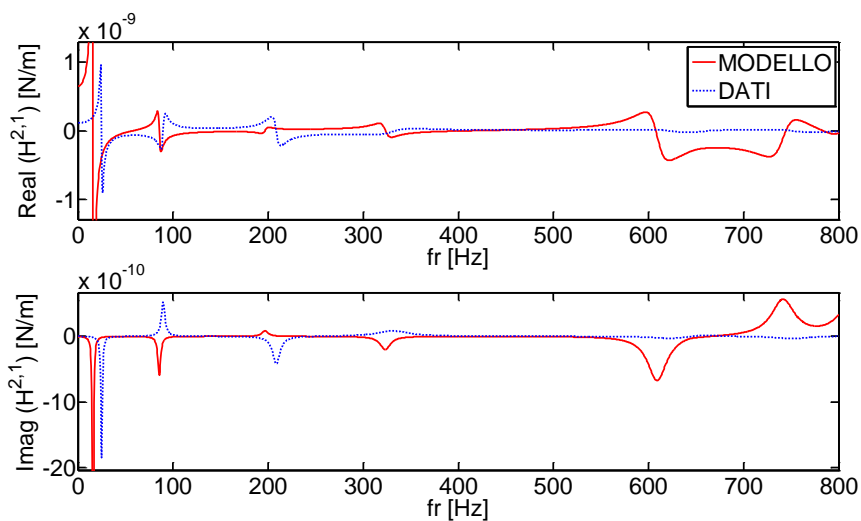


Figura 3. Confronto delle FRF (ingresso gdl 1, uscita gdl 2) prima della procedura di updating: dati di ingresso (linea blu punteggiata) e modello (linea rossa continua).

Risultati della procedura di updating

I parametri coinvolti nella procedura di updating sono i coefficienti del modello di vincolo κ_{ij}^{33} ($m^{33}=7$, $p^{33}=2$, $n^{33}=1$, $q^{33}=0$ in Eqn.(9)) e quelli del modello di smorzamento γ_z (funzioni B-spline quadratiche con $n^z=7$, $f_{ST}=0$ Hz e $f_F=1.2 \cdot 800=960$ Hz).

La procedura è inizializzata ponendo tutti i coefficienti $\kappa_{ij}^{33} = 1 \cdot 10^{11}$ e tutti i coefficienti di smorzamento $\gamma_z = 0.01$. Tutte le sei FRF sono state considerate. In Fig.3-4 sono presentate alcune FRF utilizzate come ingresso della procedura di updating, con le corrispondenti FRF generate dal modello con i parametri di partenza.

La Fig.5 permette di paragonare i fattori di smorzamento modali utilizzati per simulare le misure e quelli identificati tramite la procedura proposta. Il confronto tra le FRF utilizzate come ingresso e quelle ottenute dopo la procedura di updating sono visibili nelle Fig.6-7. La Tabella 1 presenta le prime nove frequenze naturali del modello utilizzato per generare le FRF di input, e del modello prima e dopo la procedura di updating.

Tabella 1. Confronto delle prime nove frequenze naturali del modello utilizzato per generare le FRF di input (f_{MIS}), e del modello prima (f_{IN}) e dopo (f_{FIN}) la procedura di updating.

f_{MIS} Hz	f_{IN} Hz	f_{FIN} Hz
24.91	15.82	26.64
89.58	85.56	89.72
208.66	196.87	211.13
326.68	323.20	326.55
332.74	323.67	333.71
625.24	609.24	621.41
657.54	642.69	662.19
751.36	741.62	755.71
818.43	810.12	816.24

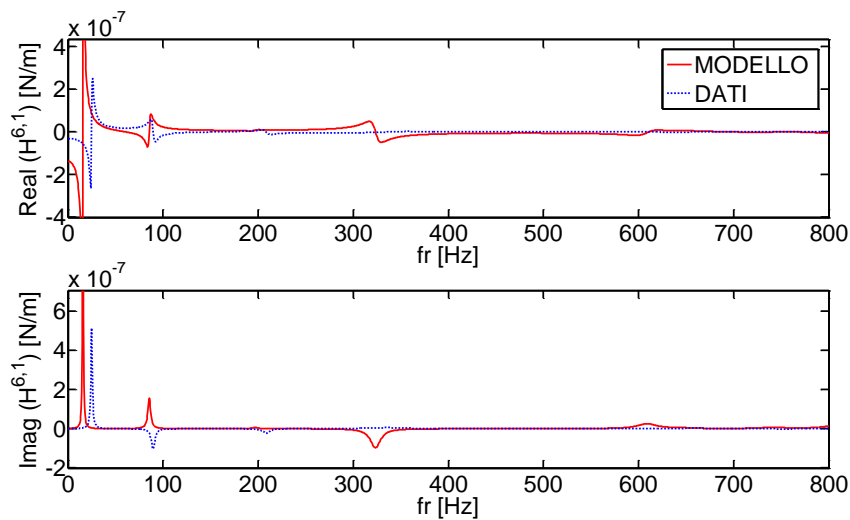


Figura 4. Confronto delle FRF (ingresso gdl 1, uscita gdl 6) prima della procedura di updating: dati di ingresso (linea blu punteggiata) e modello (linea rossa continua).

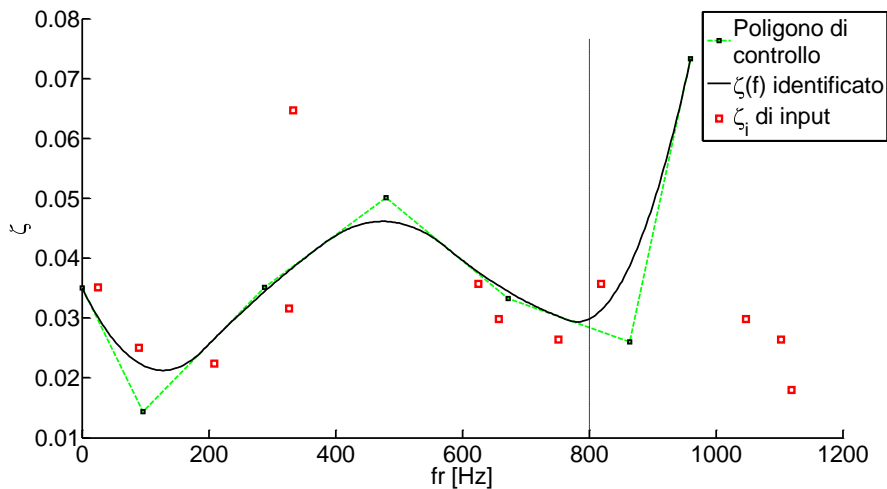


Figura 5. Confronto dei fattori di smorzamento modali ζ_i (quadrati in rosso) utilizzati per generare le FRF di input (nell'intervallo [0,800] Hz) e funzione $\zeta(f)$ identificata (linea continua nera; la linea tratteggiata verde è il poligono formato dai pdc). Il modello di smorzamento è definito sull'intervallo [0,960] Hz.

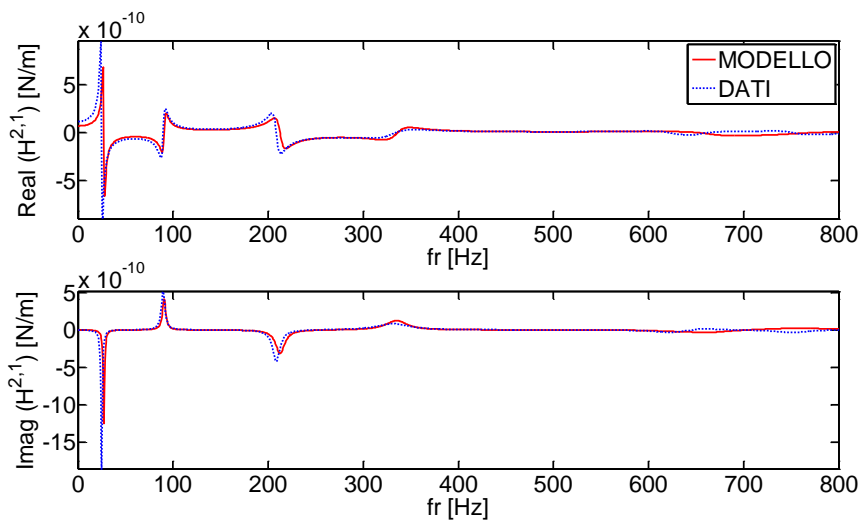


Figura 6. Confronto delle FRF (ingresso gdl 1, uscita gdl 2) dopo la procedura di updating: dati di ingresso (linea blu punteggiata) e modello (linea rossa continua).

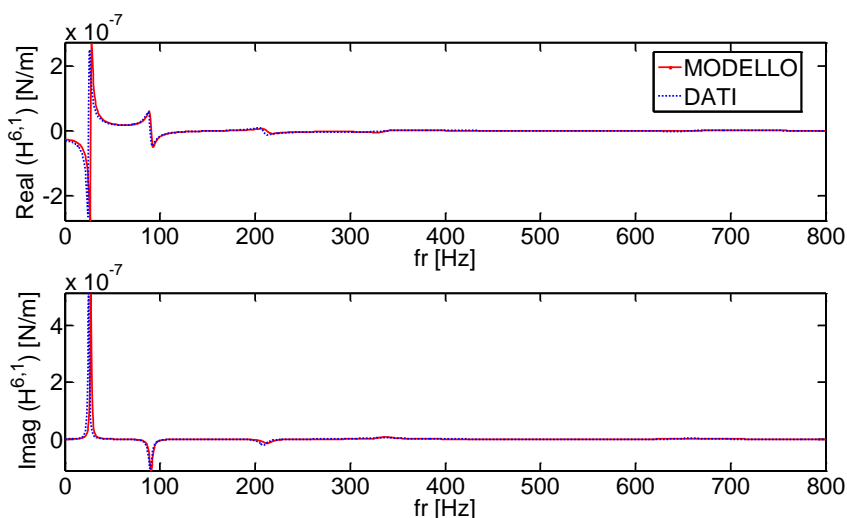


Figura 7. Confronto delle FRF (ingresso gdl 1, uscita gdl 6) dopo la procedura di updating: dati di ingresso (linea blu punteggiata) e modello (linea rossa continua).

5. CONCLUSIONI

È stata presentata una procedura di updating di un modello ad EF basato su funzioni di forma di tipo B-spline utilizzando FRF come dati di ingresso. La procedura prevede la minimizzazione la somma dei quadrati dei residui considerati a varie frequenze, e ha condotto ad un sistema di equazioni sovra-determinato che è stato risolto ai minimi quadrati. I parametri ottimizzati sono i coefficienti dei modelli di vincolo distribuito e di smorzamento proposti. Entrambi i modelli sono stati descritti mediante l'utilizzazione di funzioni B-spline. L'incompatibilità tra i gdl di misura e quelli del modello è stata risolta utilizzando le stesse funzioni di forma del modello ad EF, con un piccolo costo computazionale aggiuntivo. È stato presentato un esempio di applicazione della procedura di updating riguardante una piastra quadrata sottile incastrata su un lato. Le quattro FRF usate come ingresso della procedura sono state simulate utilizzando lo stesso modello di piastra ma utilizzando un diverso modello di vincolo e di smorzamento. I risultati numerici mostrano una buona corrispondenza con le FRF di ingresso. Le parametrizzazioni proposte permettono un efficiente adattamento dei modelli di vincolo di smorzamento al sistema da rappresentare indipendentemente dal numero di gdl del modello FE. È necessario verificare l'approccio usando dati realmente misurati come ingresso, al fine di esaminare l'effetto degli errori di modello e di misura. Ricerche future saranno dirette sulla scelta ottimale delle frequenze da considerare nella procedura di updating.

Ringraziamenti

Il presente studio è stato realizzato all'interno del CIRI-MAM, con il contributo della Regione Emilia-Romagna, Progetto Tecnopoli.

RIFERIMENTI

- [1] Hughes, T.J.R., Cottrell, J.A. and Bazilevs, Y., 2005. "Isogeometric analysis: CAD, finite elements, NURBS, exact geometry, and mesh refinement". *Computer Methods in Applied Mechanics and Engineering*, **194**, pp. 4135–4195.
- [2] Kagan, P., Fischer, A., and Bar-Yoseph, P. Z., 2003. "Mechanically based models: Adaptive refinement for B-spline finite element". *Int. J. Numer. Meth. Engng.*, **57**, pp. 1145–1175
- [3] Debabrata, R., 2003. "c-Type method of unified CAMG and FEA. Part 1: Beam and arch mega-elements—3D linear and 2D non-linear". *Int. J. Numer. Meth. Engng.*, **58**, pp.1297–1320.
- [4] Hughes, T.J.R., Reali, A., and Sangalli, G., 2009. "Isogeometric methods in structural dynamics and wave propagation". In Proceedings of COMPDYN 2009-Computational Methods in Structural Dynamics and Earthquake Engineering, Rhodes, Greece, 22-24 June 2009.
- [5] Friswell, M. I., and Mottershead, J. E., 1995. *Finite element modal updating in structural dynamics*, Kluwer Academic Publisher, ISBN 0-7923-3431-0, Dordrecht, Netherlands.
- [6] Friswell, M.I., Mottershead, J.E., and Ahmadian, H, 2001. "Finite-Element Model Updating Using Experimental Test Data: Parametrization and Regularization". *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 359, 1778, Experimental Modal Analysis (Jan. 2001), pp. 169-186.
- [7] D'ambrogio W., and Fregolent A., 2000. "Robust dynamic model updating using point antiresonances". Proceedings of the 18th International Modal Analysis Conference, pp. 1503-1512, San Antonio, Texas.
- [8] Esfandiari, A., Bakhtiari-Nejad, F., Rahai, A., and Sanayei, M., 2009. "Structural model updating using frequency response function and quasi-linear sensitivity equation". *Journal of Sound and Vibration*, Vol. **326**, 3-5, pp. 557-573, ISSN 0022-460X.
- [9] Lin, R.M., and Zhu, J., 2006. "Model updating of damped structures using FRF data". *Mechanical Systems and Signal Processing*, **20**, pp. 2200-2218.
- [10] Gabriele S., Valente, C., and Brancaloni, F., 2009. "Model calibration by interval analysis". Proceedings of XIX AIMETA CONFERENCE, Ancona, Italy, September 14-17, 2009.
- [11] Kye-Si Kwon¹, Rong-Ming Lin , 2004. "Frequency selection method for FRF-based model updating". *Journal of Sound and Vibration*, **278**, pp. 285–306
- [12] Carminelli, A., Catania, G., 2010. "B-Spline Shell Finite Element Updating By Means Of Vibration Measurements". Participation in the book *Advances in Vibration Analysis Research*, INTECH, ISBN 978-953-307-209-8.
- [13] Piegl, L., and Tiller, W., 1997. *The NURBS Book, 2nd Edition*. Springer-Verlag, ISBN 3-540-61545-8, New York, NY, USA.
- [14] Carminelli, A. and Catania, G., 2007. "Free vibration analysis of double curvature thin walled structures by a B-spline finite element approach". Proceedings of ASME IMECE 2007, pp. 1-7, Seattle (Washington), USA, 11-15 November 2007.
- [15] Cook, R.D., Malkus, D.S., Plesha, M.E., and Witt, R.J., 1989. *Concepts and applications of finite element analysis*, J. Wiley & Sons, ISBN 0-471-35605-0, New York, NY, USA.

EFFETTO DEL CONTENUTO DI IDROGENO E DELLE VELOCITA' DI RAFFREDDAMENTO SULLA MICROPOROSITA' NELLA PRODUZIONE PER FUSIONE DI CAMPIONI A DIFETTOLOGIA CONTROLLATA PER PROVE DI FATICA

Ivan Todaro
DIEM
Università di Bologna
E-mail: ivan.todaro@mail.ing.unibo.it

Rosario Squatrito
DIEM
Università di Bologna
E-mail: Rosario.Squatrito@unibo.it

Lorella Ceschini
SMETEC
Università di Bologna
E-mail: ceschini@bomet.fci.unibo.it

Andrea Morri
SMETEC
Università di Bologna
E-mail: a.morri@unibo.it

Luca Tomesani
DIEM
Università di Bologna
E-mail: luca.tomesani@unibo.it

Sommario. La produzione di un numero statisticamente rilevante di provini consistentemente simili in termini di microstruttura e difettologia da utilizzare in prove di fatica, richiede un controllo molto spinto della conduzione delle attrezzature di colata e delle condizioni di processo. In questo lavoro si presenta un esperimento per la produzione, tramite processi di colata in conchiglia per gravità, di provini in lega di alluminio a caratteristiche microstrutturali e difettologiche controllate. I principali parametri di processo sottoposti a controllo sono stati le velocità di raffreddamento imposte al getto e la quantità di idrogeno nel bagno fusorio. Si è inoltre

tentato un controllo della quantità di ossidi nel bagno, che agendo come siti di nucleazione delle porosità, possono influire sul loro numero e dimensione.

Si è valutata la distribuzione di alcune grandezze rilevanti nei getti prodotti (SDAS, Area % delle Porosità ecc.) allo scopo di validare i criteri di progettazione dello stampo e della conduzione dell'esperimento.

Parole Chiave: colata in conchiglia, leghe di alluminio, provini, fatic, microporosità

1. INTRODUZIONE

Le leghe di Alluminio della serie A356/A357 trovano largo impiego nel settore automotive grazie ad eccellenti caratteristiche di colabilità, resistenza a corrosione e specialmente all'ottimo rapporto fra caratteristiche meccaniche e densità. I processi di formatura da liquido introducono inevitabilmente difetti peculiari che possono abbattere le caratteristiche meccaniche delle leghe in termini di allungamento percentuale e soprattutto di vita a fatica.

I difetti che maggiormente incidono sulle prestazioni a fatica di un getto sono le porosità da ritiro e da gas [1]. Solo quando le porosità sono trascurabili, gli effetti di altri difetti divengono apprezzabili [2-3]

Per valutare l'effetto di tali difetti ed altre caratteristiche microstrutturali sulla resistenza a fatica, è fondamentale produrre provini a difettologia e microstruttura controllate. A tale scopo è necessario svolgere uno studio approfondito sull'attrezzatura di produzione e sulla conduzione di tutto il processo di colata, a partire dai protocolli di trattamento della lega liquida, di riempimento dello stampo, fino alle modalità di condizionamento del raffreddamento del getto. In tutto lo studio, la simulazione numerica del processo riveste un ruolo fondamentale.

2. DESCRIZIONE DELL'ESPERIMENTO

Gli scopi della sperimentazione sono stati:

1. Studiare una geometria semplice per i getti in modo da rendere possibile, per ciascuno di essi, l'estrazione di set di campioni sostanzialmente identici in termini di microstruttura e distribuzione dei difetti
2. Differenziare i getti fra di loro per microstruttura e porosità
3. Produrre un gradiente di microstruttura e porosità nella sezione di ciascun getto

Progettazione della conchiglia

I requisiti generali per la progettazione della conchiglia sono stati:

1. Possibilità di variare le velocità di raffreddamento per ciascuna colata
2. Forzare la direzione del Flusso di Calore durante la solidificazione per indurre superfici isoterme planari, in modo da ottenere un gradiente di microstruttura nella sezione del getto
3. Condizioni di riempimento tali da evitare turbolenze ed intrappolamento di ossidi

Questi requisiti si sono concretizzati nella geometria presentata in Fig.1. Si tratta di una conchiglia conforme alla pratica della fusione in conchiglia per gravità, basata su un bastone di colata convergente, una curva di 90° ad ampio raggio di raccordo, seguita da un condotto orizzontale ed un attacco di colata singolo nella parte inferiore del getto in modo da ottenere un riempimento "in sorgente".

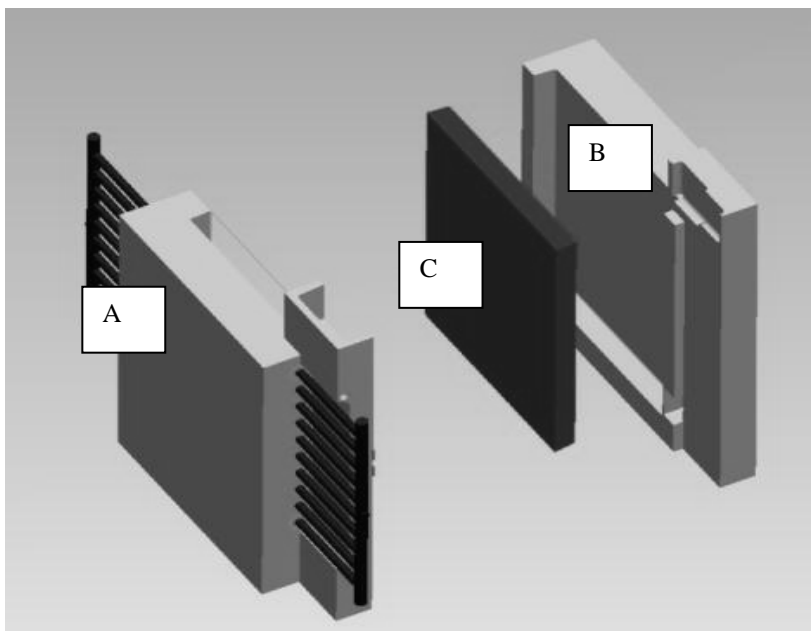


Figura 1. Disegno della conchiglia

Il getto è una semplice lastra verticale di spessore 30 mm, 300 mm in altezza e 250 mm in larghezza.

La conchiglia è composta da due elementi (A e B in Fig.1) ricavati per asportazione di truciolo a partire da due blocchi in acciaio da costruzione C40.

Per controllare le velocità di raffreddamento dei getti l'elemento denominato (A), è stato dotato di un sistema di canali posizionati a 20 mm dall'interfaccia getto/stampo, alimentabili con portate note di aria o acqua. L'elemento (A) è inoltre riscaldabile attraverso torce a metano.

Una lastra di materiale refrattario spessa 30 mm (C), è stata alloggiata in una tasca ricavata nell'elemento (B) per inibire il raffreddamento del getto da una delle due superfici maggiori, allo scopo di ottenere un gradiente delle velocità di raffreddamento nel getto.

Per valutare la distribuzione e l'evoluzione delle temperature nella conchiglia, otto termocoppie di tipo "K" da 0.5 mm di diametro sono state disposte a 5 mm e 10 mm dalla superficie getto/stampo.

Per definire le sezioni notevoli del sistema di alimentazione, è stata imposta una velocità massima di 0.5 m/s, pari alla velocità critica indicata in letteratura come limite per evitare difetti tipici del riempimento di un getto [4-5]

Il dimensionamento dei canali di colata è stato attentamente studiato per mezzo di simulazioni numeriche utilizzando il software commerciale PROCast V.2008 allo scopo di garantire un flusso per quanto possibile privo di turbolenze, evitare ingressi prematuri del fuso nella figura e garantire le giuste velocità di riempimento. Per ottenere tali scopi il sistema di colata è stato dotato di:

1. Uno swirler cilindrico, posizionato sopra al filtro (disposto orizzontalmente) per assorbire picchi di pressione dovuti al transitorio di riempimento della tazza e del bastone di colata
2. Una trappola alla fine del canale di alimentazione orizzontale, per evitare l'ingresso nella figura degli ossidi che si formano sul fronte di avanzamento del metallo

Le prime analisi numeriche di riempimento della conchiglia hanno mostrato un ampio ricircolo nel piano verticale del getto che potrebbe forzare l'innesco di moti convettivi del fuso durante la solidificazione. Per limitare tale effetto si è proceduto al dimensionamento degli attacchi di colata cercando di minimizzare il momento della quantità di moto del ricircolo. Campagne successive di simulazione hanno dimostrato come il rapporto fra il primo tempo di solidificazione (sotto la T_{liq}) ed il tempo di insorgenza del flusso convettivo implicassero limitate differenze di temperatura nel getto alla fine del riempimento.

L'analisi numerica della solidificazione (Fig.2) ha mostrato una stratificazione delle isoterme di raffreddamento un gradiente delle velocità di solidificazione, il cui fronte planare procede dalla zona "fredda" del getto verso quella a contatto con la lastra di materiale refrattario. Tale osservazione conforta la possibilità di estrarre da ogni getto almeno un set di campioni caratterizzati da condizioni di raffreddamento pressoché identiche.

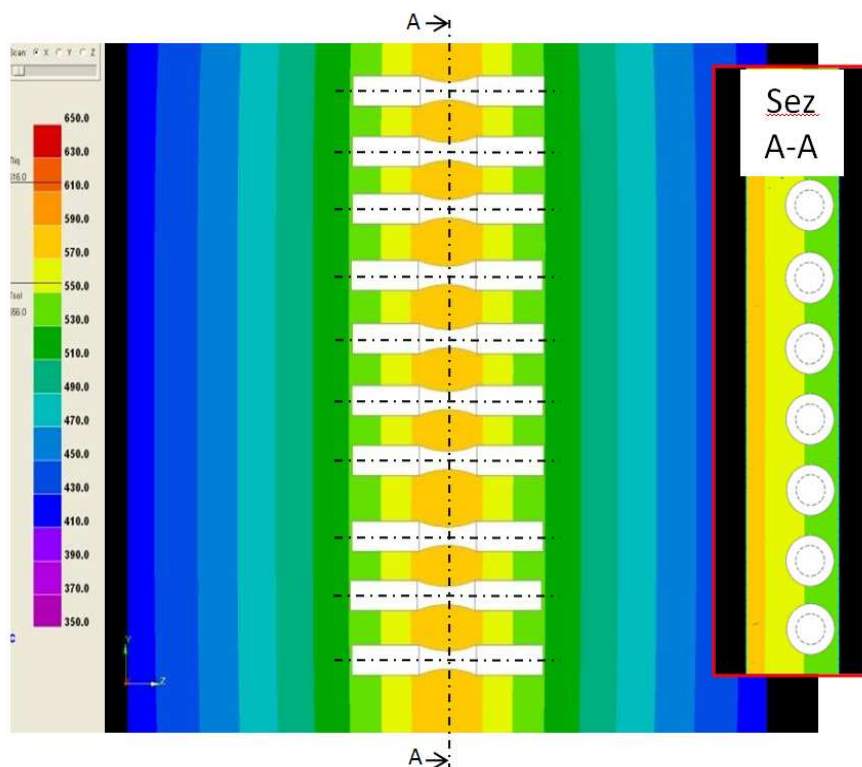


Figura 2. Analisi numerica della solidificazione: è possibile notare una stratificazione planare delle isoterme di raffreddamento nella sezione longitudinale del getto

Trattamento del bagno fusorio

Per differenziare i livelli di microporosità si può agire sulle modalità di alimentazione del getto (per promuovere o inibire la creazione di porosità da ritiro) o sulla quantità di idrogeno disciolto nel bagno di fusione, onde generare porosità da gas. Dato che il controllo delle porosità da ritiro non è agevole da realizzare ed è fortemente condizionato da fattori geometrici, si è scelto come “driver” delle porosità il tenore di idrogeno nel bagno di fusione.

Per ottenere livelli diversificati e controllati di idrogeno in soluzione nel bagno fusorio si sono messe a punto apposite attrezzature ed un opportuno disciplinare delle attività di trattamento del bagno di fusione.

Per partire sempre dalle medesime condizioni si è scelto di sottoporre ognuna delle cariche di lega fusa nel forno di attesa, allo stesso trattamento preliminare consistente di:

1. Un attento processo di scorifica
2. Uno o più cicli di degasaggio mediante lancia rotante ad argon (Fig. 3)
3. Misurazione, mediante sonda Foseco Alspek-H (Fig. 3), del livello di idrogeno nel bagno. Il livello target è di 0.08-0.10 ml H₂ /100g lega

A partire da tale trattamento preliminare è stato possibile raggiungere i livelli di idrogeno desiderati (Tab.1) facendo gorgogliare nel bagno una miscela al 10% di Idrogeno in Argon; come apparato di gorgogliamento è stata utilizzata la lancia di degasaggio (Fig. 3).



Figura 3. Lancia rotante e sonda per la misurazione della quantità di idrogeno nel bagno (Foseco Alspek-H)

Tabella 1. Livelli di Idrogeno

Livello di Idrogeno (HL)	HL0 – Molto Basso	HL1 - Basso	HL2 -Medio	HL3-Alto
Val. (mLH ₂ /100g)	0.08	0.13-0.16	0.23-0.24	0.30-0.33

Tabella 2. Livelli di ossidazione

Livelli di ossidazione (OL)	OL0	OL1	OL2	OL3
Val (minuti)	0	1	3	5

Un ulteriore scopo della sperimentazione è stato studiare l'effetto della quantità di ossidi presenti nel bagno sul processo di nucleazione delle porosità. Si è implementato a tale scopo un sistema di insufflaggio di quantità crescenti di aria nel bagno. A causa dell'ossigeno presente nell'aria insufflata, il gorgogliamento provoca la parziale ossidazione dell'alluminio e l'intimo rimescolamento degli ossidi in seno al bagno fusorio. Il livello di ossidazione è stato determinato in funzione del tempo di insufflaggio (Tab.2).

Analisi microstrutturale e difettologica

Per valutare la corrispondenza fra gli obiettivi prefissi e la realtà sperimentale sono state svolte una serie di misurazioni metallografiche in corrispondenza della zona di estrazione dei provini.

Le analisi metallografiche sono state eseguite nella zona centrale del Campione Metallografico 0 indicato in Fig. 5, per ognuno dei getti prodotti.

Le porzioni di materiale sono state estratte, inglobate e preparate per l'osservazione secondo le consuete tecniche di preparazione metallografica [6]. L'osservazione della microstruttura e della microporosità è stata svolta al microscopio ottico a riflessione e le immagini acquisite sono state elaborate mediante l'utilizzo di software per l'analisi di immagine. I dati acquisiti sono stati poi elaborati statisticamente mediante l'ausilio di fogli di calcolo commerciali per la valutazione di:

1. Distanza fra i bracci secondari delle dendriti (SDAS [μm])
2. Area Percentuale dei difetti (DAF%)
3. Numero dei difetti per cm^{-2} (N_d [cm^{-2}])
4. Massimo diametro di Feret (dimensione maggiore del rettangolo che inscrive il difetto [μm])
5. "Roundness" R, definito come:

$$R = \frac{(\text{Perimeter}_{\text{defect}})^2}{4\pi \text{Area}_{\text{defect}}} \quad (1)$$

Dove $R=1$ per pori perfettamente circolari, mentre valori maggiori di 1 indicano morfologie dei difetti più complesse.

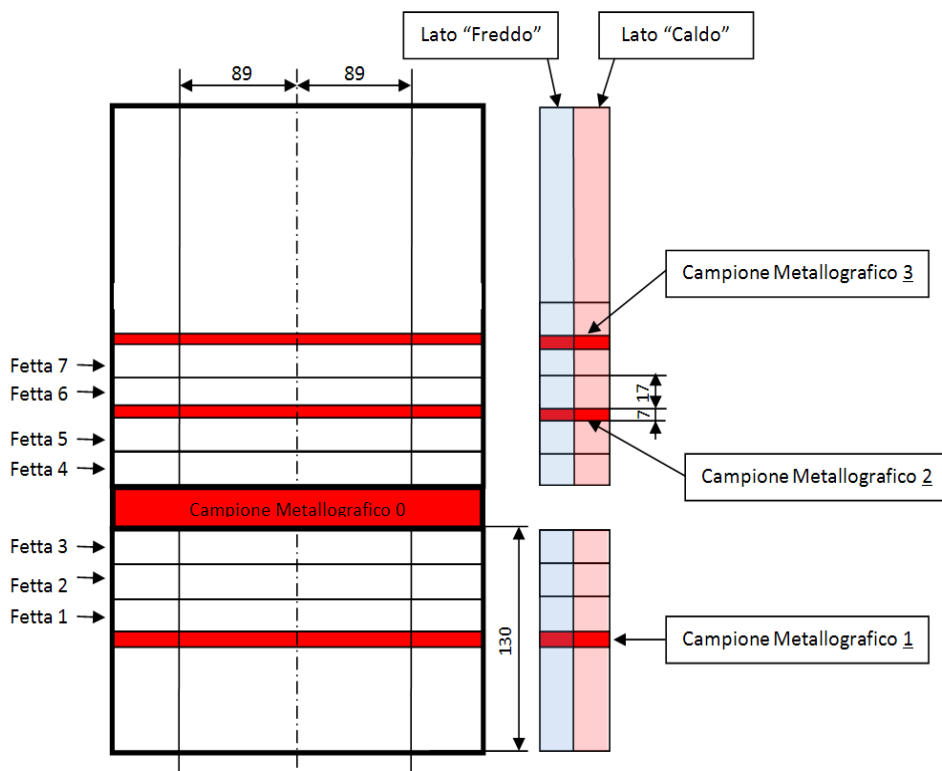


Figura 4. Zona di estrazione del campione metallografico sottoposto ad analisi

3. RISULTATI E DISCUSSIONE

Per l'analisi dello SDAS si presentano i risultati relativi ad alcune piastre caratterizzate dalla stessa temperatura di partenza della conchiglia, in quanto rappresentative dell'intera esperienza.

Come è possibile notare da Tab. 3 e da Fig. 5 esiste una chiara tendenza dello SDAS spostandosi dal lato del getto a contatto con la conchiglia (il lato da cui l'estrazione del calore è stata più vigorosa) verso il lato "caldo" del getto, quello cioè a contatto con la lastra di materiale coibente. La variazione di SDAS risulta inoltre sufficientemente consistente, essendo lo scarto medio di ogni misurazione rispetto alla media misurata nello stesso punto per tutti i getti, raccolto in un intervallo di 5 μm al massimo.

La produzione di porosità (Tab. 4 e Fig. 6) è estremamente sensibile al contenuto iniziale di idrogeno nella lega. Si riesce quindi ad attingere valori medi di porosità sufficientemente differenti, agendo su questo parametro. È inoltre possibile notare una variazione dell'entità della porosità percentuale nella sezione, spostandosi dalla parte del getto a contatto con la conchiglia verso la parte a contatto col materiale coibente, concordemente con la diminuzione delle velocità di raffreddamento del getto.

Tabella 3. Misurazioni puntuali dello SDAS per quattro dei getti prodotti: valori puntuali nella sezione

Distanza dalla conchiglia (Lato Freddo) [mm]	Getto 1 (SDAS) [mm]	Getto 2 (SDAS) [mm]	Getto 3 (SDAS) [mm]	Getto 4 (SDAS) [mm]
1.9	41	39	38	38
5.9	44	43	42	41
9.9	46	48	42	43
13.9	50	49	46	48
17.9	51	53	49	47
21.9	51	55	54	51
25.9	53	57	54	53
29	50	50	47	47

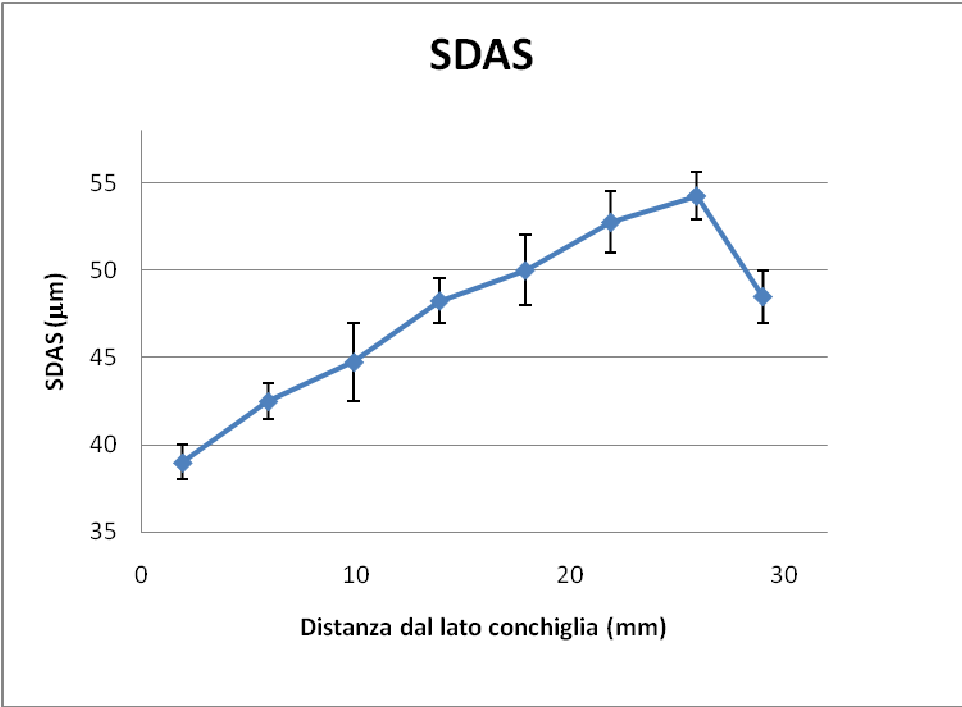


Figura 5. Andamento medio dello SDAS per le quattro lastre selezionate nella sezione sottoposta ad indagini metallografiche.

Tabella 4. Porosità medie nelle famiglie di getti a contenuto iniziale di idrogeno differenziato: valori puntuali nella sezione

	Idrogenazione HL0 [ml/100g]	Idrogenazione HL1 [ml/100g]	Idrogenazione HL2 [ml/100g]	Idrogenazione HL3 [ml/100g]
Distanza dalla conchiglia (Lato Freddo) [mm]	Porosità %	Porosità %	Porosità %	Porosità %
1.90	0.03	0.04	0.52	1.26
5.90	0.00	0.06	0.53	1.45
9.90	0.00	0.20	0.50	1.25
13.90	0.00	0.27	0.57	1.14
17.90	0.00	0.39	0.71	1.61
21.90	0.07	0.32	0.82	1.88
25.90	0.01	0.79	0.94	1.63
29.00	0.16	0.70	0.73	1.87

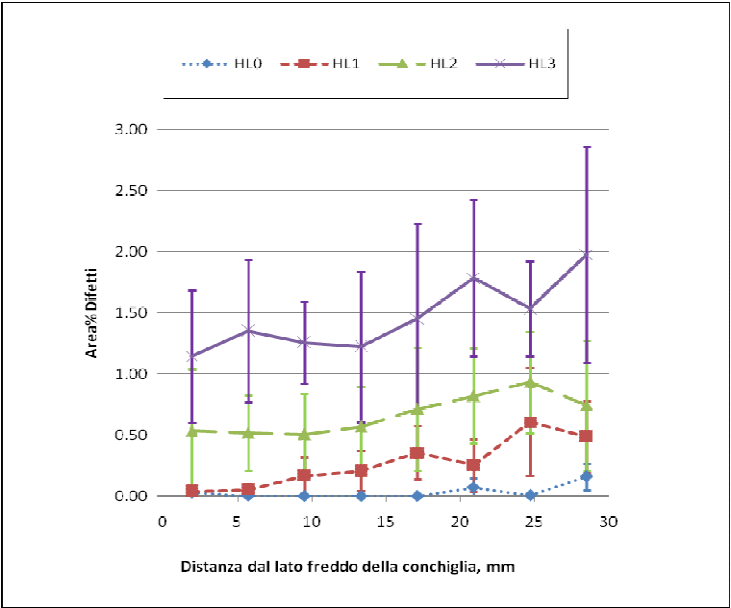


Figura 6. Comparazione degli andamenti delle porosità nelle famiglie di getti a contenuto iniziale di idrogeno differenziato, nella sezione sottoposta ad indagini metallografiche.

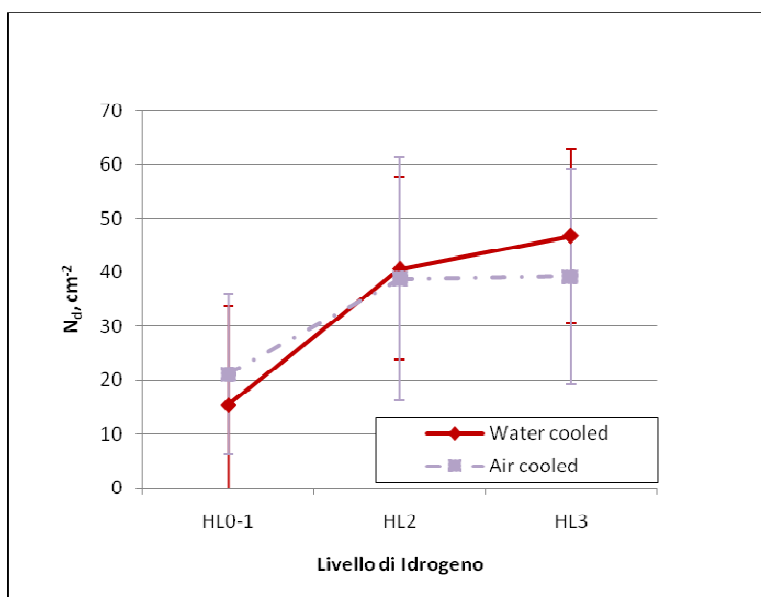


Figura 7. Effetto del contenuto di idrogeno sul numero di difetti per unità di superficie

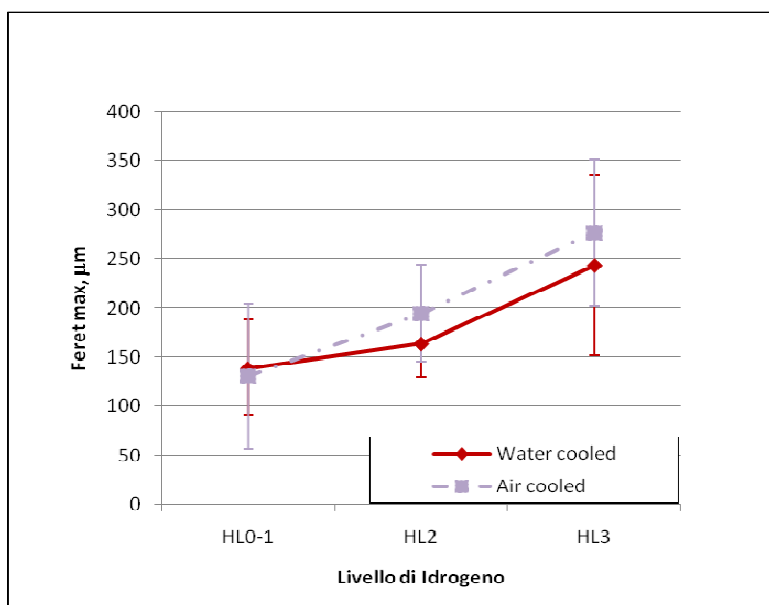


Figura 8. Effetto del contenuto di idrogeno sulla dimensione dei difetti espressa in termini di Feret massimo

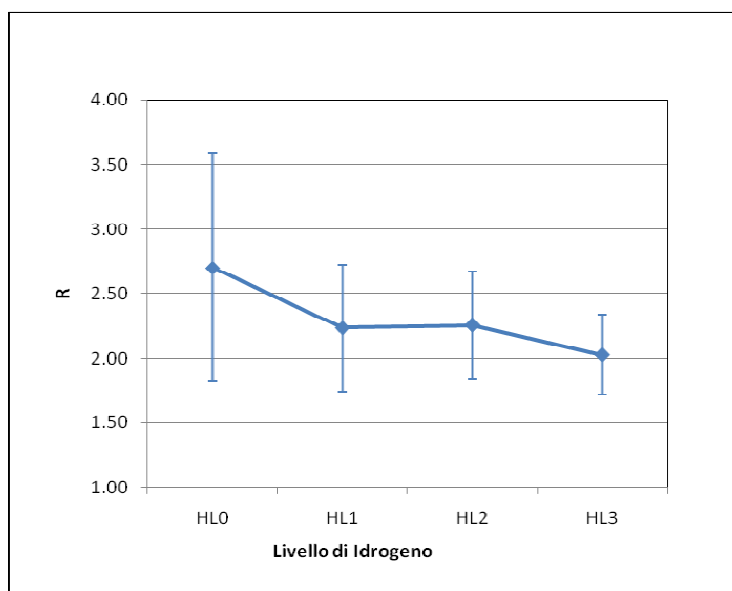


Figura 9. Effetto del contenuto di idrogeno sulla forma dei difetti, espresso in termini di Roundness

Il numero di difetti per unità di superficie tende ad aumentare con il livello di idrogeno nel bagno (Fig.7). Tale tendenza rimane valida indipendentemente dal tipo di raffreddamento della conchiglia durante la solidificazione del getto.

Concordemente alle attese, il livello di idrogeno ha un impatto anche sulle dimensioni medie dei difetti, espresse in termini di Feret massimo (Fig. 8). I getti raffreddati ad aria inoltre presentano dimensioni delle porosità maggiori (circa 40 μm) rispetto a quelli raffreddati ad acqua: ciò conferma l'influenza delle velocità di raffreddamento sul processo di accrescimento dei difetti.

Come si nota da Fig. 9, la quantità di idrogeno nel bagno influenza anche la morfologia dei pori: all'aumentare della quantità di gas disciolto corrispondono valori di Roundness inferiori, a testimonianza di una transizione verso forme più arrotondate dei difetti.

All'attuale stato degli studi, l'analisi delle porosità non ha evidenziato un chiaro impatto del livello di ossidazione del fuso sulle caratteristiche dei difetti. L'osservazione diretta delle attività sperimentali ha dimostrato come l'insufflaggio di aria in bolle estremamente fini e disperse, tipica del processo di rotary degassing, ha sortito l'effetto di abbattere il tenore di idrogeno del bagno e promuovere la risalita in superficie, per flocculazione inversa, degli ossidi generati. In sostanza è possibile affermare che l'insufflaggio di aria tramite lancia rotante non è un metodo efficiente per creare ossidi che rimangano dispersi nel bagno. Non è quindi stato possibile studiare in maniera rigorosa l'effetto dell'aggiunta di fasi nucleanti per le porosità nel bagno.

4. CONCLUSIONI

In questo lavoro è stata proposta una procedura sperimentale il cui scopo è quello di creare materiale a distribuzione nota e controllata di porosità e caratteristiche microstrutturali, per la produzione di provini a fatica.

La geometria scelta per l'attrezzatura e le modalità di conduzione dell'esperimento hanno consentito di generare famiglie di getti da cui estrarre provini per prove di caratterizzazione a fatica con una distribuzione di microstruttura e difettosità adeguatamente controllate.

Il controllo dell'ossidazione del bagno per lo studio degli effetti sulla nucleazione delle porosità si è dimostrato invece di difficile esecuzione.

Ringraziamenti

Si desidera ringraziare: Ferrari SpA, nella persona dell'Ing. Gianluca Pivetti, per aver concretizzato la possibilità di eseguire le attività sperimentali descritte; L'Ing. Lorenzo Pivetti e i Sigg. Stefano Calcaterra e Giuseppe Arcadipane, di Fonderia Scacchetti, per aver fornito l'ospitalità per l'esecuzione delle prove e la competenza per la conduzione delle stesse; FOSECO Italia per aver fornito le attrezzature di misura; Franco Iorio (Modelleria CPC) per la produzione della conchiglia.

BIBLIOGRAFIA

- [1] Wang QG, Apelian D, Lados DA., 2001. "Fatigue behavior of A356-T6 aluminum cast alloys. Part I. Effect of casting defects." *Journal of Light Metals*; pp.73-85.
- [2] Wang QG, Apelian D, Lados DA., 2001, "Fatigue behavior of A356/357 aluminum cast alloys. Part II. Effect of microstructural constituents." *Journal of Light Metals* " pp. 85-97.
- [3] Ceschini L, Morri, A, Sambogna G., 2008. "The effect of hot isostatic pressing on the fatigue behaviour of sand-cast A356-T6 and A204-T6 aluminum alloys." *Journal of materials processing technology*; pp.231-238.
- [4] J.Campbell and R.A. Harding, 1994 "TALAT: The freezing of Castings, Lecture 3204", European Aluminum Association.
- [5] J. Campbell, 2004. "Casting Practice - the 10 rules of castings", Elsevier, Oxford,
- [6] ASTM E3-01, 2007. "Standard Practice for Preparation of Metallographic Specimens"

ANALISI TEORICA E SPERIMENTALE DELLE SOLLECITAZIONI IN ESERCIZIO SU UN ANELLO ROTANTE DI UNA MACCHINA CELLOFANATRICE

Giorgio Olmi

*Dip. di Ing. delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia,
Università degli Studi di Bologna, Italia*

E-mail: giorgio.olmi@unibo.it

Vincenzo Dal Re

*Dip. di Ing. delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia,
Università degli Studi di Bologna, Italia*

E-mail: vincenzo.dalre@unibo.it

Sommario. *Il presente lavoro verte sulla realizzazione di un modello analitico per l'analisi strutturale di un anello rotante, presente su una macchina impacchettatrice. Dopo un breve excursus sulle caratteristiche architettoniche e dimensionali della macchina in studio, vengono illustrate la matematica del modello e i risultati raggiunti, nei termini delle reazioni vincolari iperstatiche, delle caratteristiche della sollecitazione e delle tensioni e deformazioni. Mirando ad una validazione sperimentale, gli stessi risultati analitici sono stati impiegati per pianificare prove estensimetriche, successivamente eseguite tramite la strumentazione dell'anello rotante, con accorgimenti particolari per la trasmissione del segnale. Il confronto fra i dati sperimentali ed analitici ha rilevato degli errori di pochi punti percentuali imputabili al modello, che è quindi risultato validato, aprendo le porte ad un suo uso in ambito progettuale e di definizione dei parametri di esercizio della macchina.*

Parole chiave: *anello rotante, modello strutturale analitico, estensimetria*

1. INTRODUZIONE

La pratica di fare uso di film plastici trasparenti nel confezionamento dei prodotti iniziò a diffondersi negli anni venti. Le prime applicazioni si riscontrarono negli Stati Uniti e nell'Europa Occidentale, dove i primi esemplari di macchine automatiche vennero progettati e costruiti, per operare su film plastici in Cellofan. Tali macchine erano in grado di stampare un'immagine o una scritta sull'incarto e quindi di realizzare il processo di confezionamento del prodotto. Interessanti applicazioni iniziarono quindi a farsi strada nel settore alimentare ed in quello del tabacco, nella realizzazione di incarti, che prendessero la forma del prodotto, nel caso del pacchetto di sigarette o dei piccoli snack. Successivamente, si poneva il problema di realizzare l'imballaggio per il trasporto dei prodotti verso il magazzino o il centro di distribuzione. Allo scopo, la stecca di sigarette o il pallet, nel caso di altri prodotti, dovevano essere rivestiti di film plastico, per preservarne l'integrità.

La realizzazione di macchine che realizzassero questo processo ha permesso, a livello commerciale, il passaggio da una rete formata da piccoli negozi ad una più vasta, costituita da più ampi magazzini e centri di distribuzione. Parallelamente, si sono sviluppate macchine automatiche sempre più veloci ed affidabili, che permettessero un rapido e

rigoroso impacchettamento. Questo era necessario anche alla luce anche della principale finalità di un buon incarto, che doveva garantire un'adeguata barriera contro gli agenti batterici, l'umidità, i grassi, gli odori. A sua volta, il miglioramento degli aspetti tecnici e l'utilizzo di macchine sempre più sofisticate, richiedevano l'impiego di film dotati di adeguate caratteristiche meccanico-fisiche, quali la trasparenza, la rigidezza, la resistenza al calore, la saldabilità, la scorrevolezza sulle guide [1].

Successivamente, altri passaggi delicati sono stati legati all'introduzione di nuovi materiali per l'imballaggio, quali il polipropilene a doppia orientazione o il PVC. Questo ha richiesto in molti casi un aggiornamento delle macchine disponibili nel settore del packaging, affinché fossero rese in grado di operare con i nuovi materiali, aventi talvolta caratteristiche meccaniche differenti [1]. Oggigiorno, i materiali più utilizzati nel confezionamento sono prodotti di derivazione petrolchimica, quali il Polietilene Tereftalato (meglio noto come PET), il Polivinilcloride (PVC), il polietilene, il polipropilene, il polistirene ed il poliammide [2].

Dal punto di vista tecnico, il presente lavoro focalizza la sua attenzione sulle macchine cellofanatrici, ossia sulle macchine che risultano in grado di rivestire con un film plastico pezzi di forma oblunga (ad esempio cilindrica o di parallelepipedo), quali carichi pallettizzati, con altezza dell'ordine del mezzo metro. La tipica architettura di tali macchine è costituita da un telaio che supporta un grande anello rotante. Quest'ultimo è a sua volta solidale con il gruppo che contiene la bobina del film, oltre ai dispositivi che ne conferiscono l'adeguato tensionamento. Il pezzo da avvolgere nel film plastico è quindi posto al centro della pianta dell'anello. All'inizio delle operazioni un lembo del film viene fissato al pezzo da rivestire, successivamente, un apposito sistema di movimentazione pone in rotazione l'anello ad una velocità solitamente fino a 70-80 rpm. Il film può così essere rapidamente avvolto, operando il numero di giri richiesto, per conferire all'imballaggio le adeguate caratteristiche di resistenza meccanica, e se necessario, soprattutto nel caso dell'industria alimentare, per isolare il prodotto dagli agenti esterni. Le due configurazioni possibili possono prevedere una rotazione dell'anello attorno ad un asse verticale o attorno ad uno orizzontale, la più frequente risulta la prima ivi menzionata. In tal caso, l'architettura può variare a seconda della rigidezza del telaio, che può prevedere due o quattro montanti verticali, ed una struttura aperta o chiusa tramite traverse ad andamento laterale. In corrispondenza dei montanti sono poi realizzati gli appoggi per l'anello rotante, che quindi possono essere in numero diverso a seconda delle configurazioni. E' chiaro che un maggior numero di montanti conferisce all'intera macchina una maggiore rigidezza, inoltre la presenza di molteplici punti d'appoggio rende la struttura dell'anello fortemente iperstatica, con conseguente diminuzione delle sollecitazioni e delle deformazioni in esercizio. Sull'entità di tali frecce agisce chiaramente anche la geometria dell'anello, in particolare le sue dimensioni e la forma della sezione. Solitamente, l'anello viene ricavato saldando insieme due metà, ricavate per calandratura di tubi. Le sezioni più frequentemente utilizzate sono la classica anulare e quella quadrata o rettangolare cava, quest'ultima se adeguatamente dimensionata, porta solitamente ai migliori benefici in termini di maggiore rigidezza a parità di massa in movimento. Le macchine con maggiore rigidezza sono quelle solitamente di qualità superiore, potendo garantire un lavoro di imballaggio accurato e la massima affidabilità, anche nel caso di uso continuativo alla massima velocità consentita. E' chiaro che queste maggiori prerogative comportano costi di produzione più elevati, che in ambito di mercato si traducono in prezzi di listino più alti. La Fig. (1) mostra alcuni esempi di macchine ad asse rotante verticale ed orizzontale, con diverse caratteristiche di rigidezza.



Figura 1. Esempi di macchine avvolgitrici ad asse verticale a 4 montanti (a) [3], ad asse verticale, con due montanti con sbalzo (b, c) [3-4] ed ad asse orizzontale per imballaggi di piccole (d) e grandi dimensioni (e) [5].

Il presente lavoro si focalizza su una moderna macchina cellofanatrice, del tipo ad anello rotante attorno ad un asse verticale, i cui dettagli costruttivi e dimensionali saranno illustrati nel Par. 2. Tale macchina si colloca in una fascia alta del mercato e intende differenziarsi dalla concorrenza per la presenza di numerosi accorgimenti brevettati, che migliorano l'accuratezza dell'imballaggio, unitamente ad un'elevata rigidità, dovuta alla struttura chiusa del telaio ed al vincolamento dell'anello su otto punti, sia in direzione verticale che in direzione radiale, quindi con una forte iperstaticità. La massima accuratezza nell'esercizio deve essere solitamente abbinata alla massima affidabilità, la quale risulta legata allo stato di sollecitazione che si raggiunge nel componente più caricato: l'anello. Tale sollecitazione può variare sensibilmente nelle differenti condizioni di esercizio, a seconda della velocità di rotazione, o della massa dei gruppi trasportati, atti all'avvolgimento e tensionamento del film plastico.

Il cosiddetto "problema dell'anello caricato" è molto comune nella meccanica, con diversi esempi riportati in [6]. La ricerca della massima rigidità, con conseguente ipervincolamento, rende però la struttura presente sulla macchina in studio particolarmente unica e non è possibile riscontrare in letteratura modelli di tipo analitico, che permettano di

valutarne le sollecitazioni in funzione dei diversi parametri dimensionali e di esercizio. Il problema potrebbe chiaramente essere affrontato per via numerica, ma un modello agli elementi finiti, se da un lato rappresenterebbe l'approccio più semplice per fornire lo stato di sollecitazione corrente, sarebbe inadatto per compiere analisi comparative del tipo "What... if...", volendo cioè determinare rapidamente quanto la variazione di un parametro possa influenzare lo stato di sollecitazione. L'analisi della letteratura mostra inoltre come ben raramente si siano effettuati studi sperimentali su anelli rotanti, con acquisizione in tempo reale dei segnali deformativi, così da comprendere che tipo di oscillazioni essi subiscano alle diverse velocità o se vi siano effetti dinamici all'avviamento od all'arresto. Tali prove sono solitamente complicate dalla necessità di acquisire segnali da un corpo in movimento, il che richiede l'adozione di tecnologie wireless o con contatti striscianti.

Gli obiettivi del presente lavoro, a fronte delle motivazioni sopra riferite, sono pertanto i seguenti:

- Valutare lo stato tensionale e deformativo dell'anello per via analitica, in funzione di tutti i parametri geometrico-dimensionali e di esercizio, con particolare riferimento alla velocità di rotazione.
- Effettuare una stima sperimentale delle sollecitazioni in esercizio, considerando diverse velocità nell'abituale campo di lavoro della macchina.
- Validare il modello analitico, tramite confronto qualitativo e quantitativo degli andamenti riscontrati rispetto ai dati sperimentali.
- Prevedere la sicurezza nel caso di un eventuale aumento della velocità massima di rotazione.

2. LA MACCHINA IN STUDIO E L'ANELLO ROTANTE

Un complessivo della macchina in studio è mostrato in Fig. (2). Come si può notare, essa si sviluppa in verticale, utilizzando un telaio a quattro montanti: la struttura è ulteriormente irrigidita attraverso l'inserzione di travi ad andamento trasversale. Questo accorgimento è reso necessario anche dalle elevate dimensioni della macchina, avente quote d'ingombro di 2,4 m x 2,4 m x 4,2 m, per una massa complessiva di 3500 kg. Tali elevate dimensioni permettono di processare anche imballaggi voluminosi. In corrispondenza dei quattro montanti sono collocati i sostegni per l'anello rotante. In particolare, i vincoli sono costituiti da ruote gommate che agiscono direttamente sull'anello, entro opportune guide, per impedire spostamenti verticali e radiali.

Il dettaglio dell'anello è illustrato nella Fig. (3). Si può osservare (Fig. (3a)) come, in corrispondenza dei quattro angoli, ognuno dei vincoli sia sdoppiato e realizzato con una coppia di ruote gommate poste ad una distanza angolare di 28° l'una dall'altra. Si hanno pertanto quattro coppie di ruote gommate nella direzione radiale ed altre quattro in quella verticale, per un totale di otto ruote in ambo le direzioni. L'anello presenta un diametro (linea media) di 2640 mm e sostiene un carrello porta-bobina, in grado di accogliere un rotolo del diametro di 250 mm e dell'altezza di 500 mm, ed un gruppo pre-stiro. La massa complessiva di tali componenti è pari a 150 kg: si tratta quindi di una massa che va ad insistere, con leggera distribuzione, su un punto preciso lungo la periferia dell'anello. All'estremità diametrale opposta è collocato un contrappeso, di identica massa, che permette di bilanciare la componente centrifuga.



Figura 2. Un complessivo della macchina cellofanatrice in studio.

In alcuni modelli il contrappeso può essere sostituito da un'altra stazione contenente carrello porta-bobina e gruppo pre-stiro, analoga a quella precedentemente descritta. In tale caso, si ha il doppio vantaggio di una naturale compensazione delle forze centrifughe e della possibilità di poter svolgere contemporaneamente due bobine di film, o di poterne montare una come riserva, per farla entrare in azione, nel momento in cui la prima si esaurisce. L'anello viene mosso con l'ausilio di un motore elettrico solidale con il telaio. Questo trasmette il moto ad una cinghia, che a sua volta si va ad impegnare su una guida lungo la periferia dell'anello. Sulla cinghia stessa scorrono anche le ruote gommate, che vincolano radialmente l'anello. La Fig. (3 b-c) mostra un dettaglio della guida e della cinghia che agisce su di essa, mentre viene riportata anche una vista della sezione con gli alleggerimenti relativi alla guida laterale prima citata ed alla guida sulla faccia inferiore, in corrispondenza della quale scorrono le ruote gommate per il vincolamento verticale. La Fig. (3 c) mostra anche le quote geometriche che intervengono nel progetto ed in funzione delle quali è stato sviluppato il modello analitico.

La velocità può essere regolata fino ad un massimo di 53 rpm, che rappresenta il valore, che nel presente lavoro è stato preso a riferimento per la determinazione dei carichi centrifughi massimi.

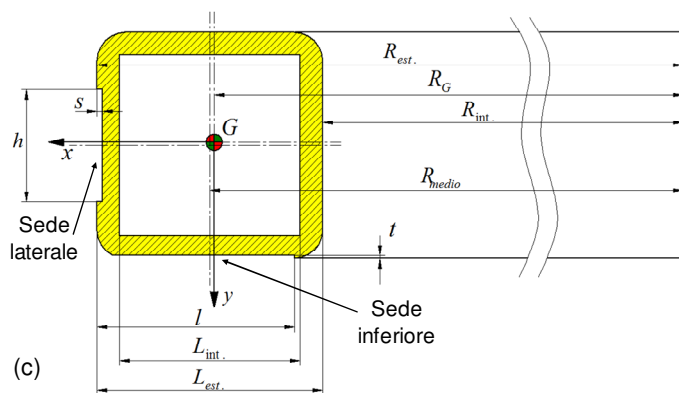
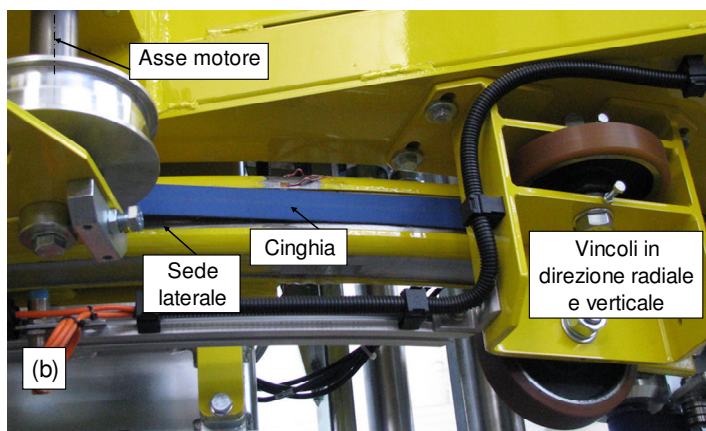
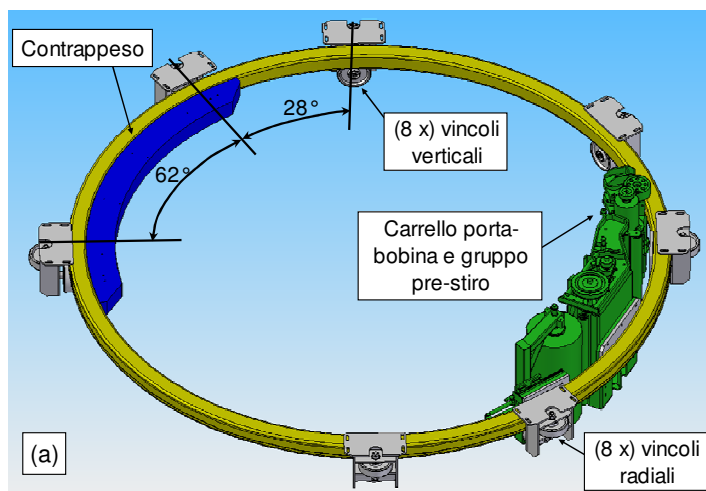


Figura 3. Dettaglio dell'anello rotante e dei componenti ivi presenti (a), particolare del sistema di movimentazione, della cinghia e dei vincoli (b) e disegno della sezione resistente con relative quote geometriche (c).

Sull'anello agiscono pertanto quattro tipologie di forze:

- Due forze centrifughe concentrate in corrispondenza del carrello porta-bobina e gruppo-pre-stiro e del contrappeso (o dell'altro carrello porta-bobina e gruppo-pre-stiro). Tali forze si possono supporre applicate in corrispondenza dei rispettivi baricentri delle masse concentrate.
- Due forze peso concentrate nei punti diametralmente opposti citati alla voce precedente.
- La forza centrifuga uniformemente distribuita dovuta alla massa propria dell'anello, realizzato in acciaio Fe510, stimata in 135 kg.
- La forza peso uniformemente distribuita, anch'essa relativa alla massa dell'anello.

3. SVILUPPO DEL MODELLO ANALITICO

Il primo passo nello sviluppo di tale modello è consistito nell'individuazione della configurazione maggiormente sollecitata, in corrispondenza della quale avviare la schematizzazione. Appare ragionevole supporre che il massimo del carico si raggiunga quando le masse concentrate si trovano al centro della campata più lunga, ossia quando l'anello è orientato come nella Fig. (3 a). In relazione alle quattro tipologie di forze evidenziate al paragrafo precedente, si è quindi proceduto nel calcolo dei relativi valori, dipendenti dalle caratteristiche geometriche e dalle condizioni di esercizio. Questo calcolo è risultato immediato nel caso delle azioni concentrate, mentre ha richiesto la valutazione di carichi equivalenti per unità di lunghezza circonferenziale nel caso delle azioni distribuite. Si rimarcano i seguenti aspetti.

- Le forze concentrate dovute all'azione centrifuga ed al peso proprio sono indicate rispettivamente con F e con P .
- Si tiene conto della forza centrifuga distribuita, mediante la pressione uniforme p , calcolata nel seguente modo (Fig. (4), Eq. (1)) [6].

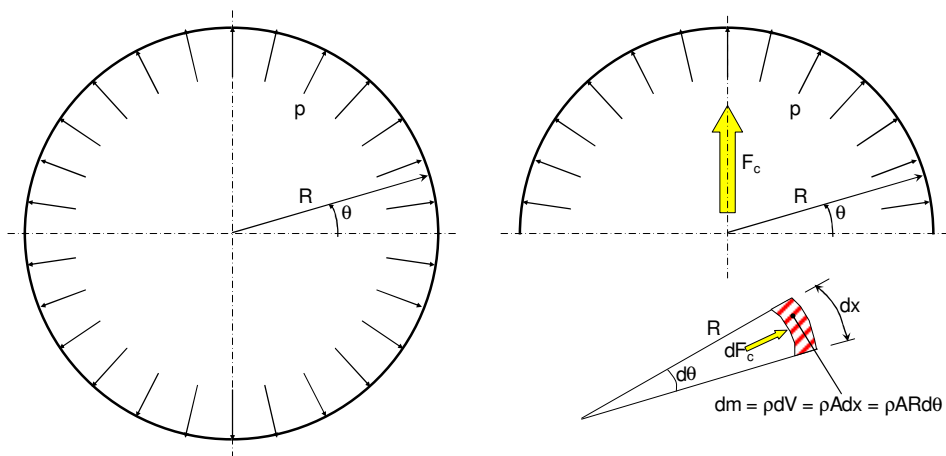


Figura 4. Determinazione della pressione p sulla superficie circonferenziale, equivalente alla forza centrifuga distribuita.

$$\begin{aligned}
dF_c &= \omega^2 R dm = \omega^2 R \rho dV = \omega^2 R \rho A dx = \omega^2 R^2 \rho A d\theta \\
F_c &= \int_0^\pi \omega^2 R^2 \rho A \sin(\theta) d\theta = 2\omega^2 R^2 \rho A \\
F_c &= 2pR \Leftrightarrow p = \frac{F_c}{2R} = \omega^2 R \rho A
\end{aligned} \tag{1}$$

Nella precedente equazione dm e dV indicano rispettivamente massa e volume dell'elemento infinitesimo, mentre A è la sezione trasversale dell'anello e ρ la massa volumica del materiale dell'anello. Si indica con F_c la risultante su metà anello della forza centrifuga distribuita.

- Infine, si è valutata una forza per unità di lunghezza per tenere conto del peso proprio dell'anello. Tale forza, indicata con q , è facilmente valutabile, dividendo il peso complessivo per la lunghezza della circonferenza media dell'anello.

Nei paragrafi che seguono verrà affrontato il problema della determinazione delle incognite iperstatiche sulla struttura dell'anello. Per fare questo, si utilizzerà il principio di sovrapposizione degli effetti, considerando separatamente le componenti radiali e verticali.

Valutazione delle reazioni sull'anello in direzione radiale

La Fig. (5) mostra la pianta dell'anello, con evidenziazione degli otto punti di vincolo e delle componenti di carico, le forze centrifughe distribuite e concentrate. Si tratta di una struttura chiusa, il cui studio risulta in prima battuta alquanto complesso; si può tuttavia osservare come vi sia una piena simmetria, geometrica e di carico, rispetto agli assi orizzontale e verticale. Proprio la presenza di questa simmetria permette di passare alle configurazioni semplificate illustrate nella stessa figura, spezzando prima l'anello in corrispondenza dell'asse orizzontale e quindi lungo quello verticale.

L'ultimo schema mostrato nell'illustrazione rappresenta la struttura effettivamente studiata per via analitica. Si nota che essa è costituita da un tratto di trave curva, corrispondente ad un quarto di circonferenza, che risulta vincolato in modo iperstatico. I vincoli considerati impediscono lo spostamento radiale in corrispondenza delle ruote gommate e gli spostamenti angolari ed assiali alle estremità. Questi ultimi vincoli nascono dall'imposizione della simmetria e dalla necessità di impedire che all'interno dell'anello si possano verificare delle compenetrazioni o dei distacchi, o che la deformata possa presentare dei punti angolosi. La struttura iperstatica è stata risolta tramite il cosiddetto "metodo delle forze", riportato ad esempio in [7]: nella stessa figura ora citata sono indicate le incognite iperstatiche X , Y e Z . Risulta al contempo evidente quale sia la struttura isostatica equivalente, per cui si sono scritte le equazioni di congruenza in Eq. (2), in funzione della coordinata angolare θ . La valutazione di tali spostamenti richiede la determinazione della linea elastica della struttura, in funzione, oltre che dalla coordinata angolare, anche di tutte le componenti di forza e momento agenti. Tale linea elastica è determinabile, risolvendo il sistema delle tre equazioni differenziali mostrato in Eq. (3): le tre componenti di spostamento che intervengono fanno riferimento allo scorrimento angolare φ , allo spostamento radiale η ed a quello in senso tangenziale u .

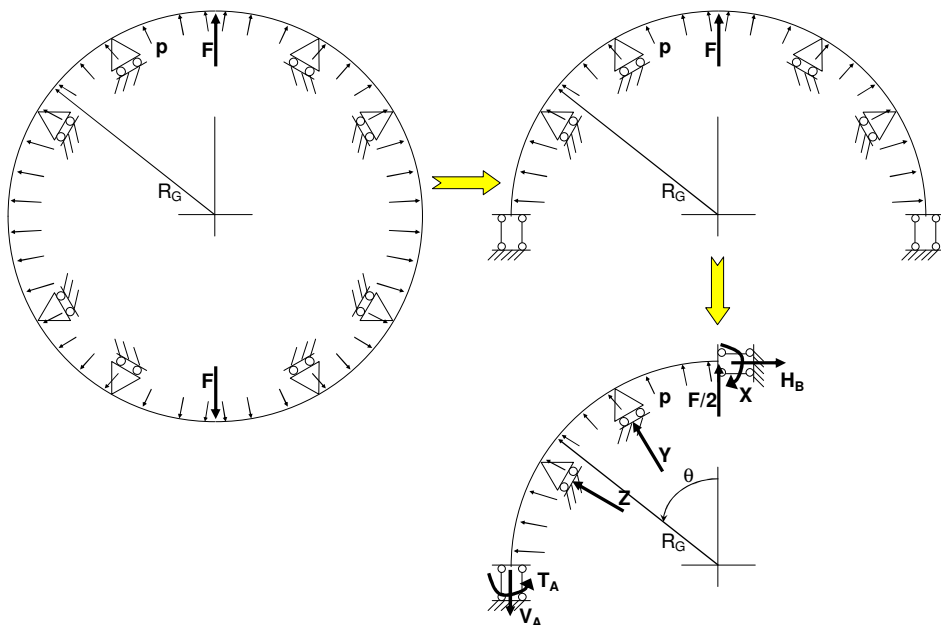


Figura 5. Schematizzazione della pianta dell'anello e determinazione per simmetria della struttura iperstatica effettivamente risolta.

$$\begin{cases} \varphi(\theta=0)=0 \\ \eta\left(\theta=\frac{28}{180}\pi\right)=0 \\ \eta\left(\theta=\frac{62}{180}\pi\right)=0 \end{cases} \quad (2)$$

$$\begin{cases} \frac{M_y(\theta)R_G}{EI_y} = -\frac{d}{d\theta}\varphi(\theta) \\ \varphi(\theta)R_G = \frac{d}{d\theta}\eta(\theta) \\ \frac{d}{d\theta}u(\theta) = \frac{N(\theta)R_G}{EA} + \eta(\theta) \end{cases} \quad (3)$$

Con E si indica il modulo elastico del materiale (assunto pari a 210 GPa), mentre con A ed I_y si indicano rispettivamente la sezione trasversale dell'anello (evidenziata in Fig. (3 c)) e il suo momento d'inerzia di area, calcolato rispetto ad un asse verticale baricentrico. Il simbolo R_G indica il raggio baricentrico dell'anello, come mostrato anche in Fig. (5). Nel sistema in Eq. (3) $M_y(\theta)$ ed $N(\theta)$ rappresentano rispettivamente il momento flettente

(positivo se tende le fibre interne) e lo sforzo normale in direzione tangenziale (positivo se di trazione) in funzione della posizione angolare θ . Questi ultimi, indicati rispettivamente con $M_y(\theta)$ e $N(\theta)$ sono così esprimibili, Eqq. (4-9):

$$M_y(\theta) = \frac{F}{2} R_G \sin(\theta) - X - \sin\left(\frac{28}{180}\pi\right) Y R_G (1 - \cos(\theta)) + \quad \text{per } 0 < \theta \leq \frac{28}{180}\pi \quad (4)$$

$$- \sin\left(\frac{62}{180}\pi\right) Z R_G (1 - \cos(\theta))$$

$$M_y(\theta) = \frac{F}{2} R_G \sin(\theta) - X + \quad \text{per } \frac{28}{180}\pi < \theta \leq \frac{62}{180}\pi \quad (5)$$

$$- \sin\left(\frac{28}{180}\pi\right) Y R_G (1 - \cos(\theta)) +$$

$$- \sin\left(\frac{62}{180}\pi\right) Z R_G (1 - \cos(\theta)) + Y R_G \sin\left(\theta - \frac{28}{180}\pi\right)$$

$$M_y(\theta) = \frac{F}{2} R_G \sin(\theta) - X - \sin\left(\frac{28}{180}\pi\right) Y R_G (1 - \cos(\theta)) + \quad \text{per } \frac{62}{180}\pi < \theta \leq \frac{\pi}{2} \quad (6)$$

$$- \sin\left(\frac{62}{180}\pi\right) Z R_G (1 - \cos(\theta)) +$$

$$+ Y R_G \sin\left(\theta - \frac{28}{180}\pi\right) + Z R_G \sin\left(\theta - \frac{62}{180}\pi\right)$$

$$N(\theta) = \frac{F}{2} \sin(\theta) + p R_G + \quad \text{per } 0 < \theta \leq \frac{28}{180}\pi \quad (7)$$

$$+ \sin\left(\frac{28}{180}\pi\right) Y \cos(\theta) + \sin\left(\frac{62}{180}\pi\right) Z \cos(\theta)$$

$$N(\theta) = \frac{F}{2} \sin(\theta) + p R_G + \sin\left(\frac{28}{180}\pi\right) Y \cos(\theta) + \quad \text{per } \frac{28}{180}\pi < \theta \leq \frac{62}{180}\pi \quad (8)$$

$$+ \sin\left(\frac{62}{180}\pi\right) Z \cos(\theta) + Y \sin\left(\theta - \frac{28}{180}\pi\right)$$

$$N(\theta) = \frac{F}{2} \sin(\theta) + p R_G + \sin\left(\frac{28}{180}\pi\right) Y \cos(\theta) + \quad \text{per } \frac{62}{180}\pi < \theta \leq \frac{\pi}{2} \quad (9)$$

$$+ \sin\left(\frac{62}{180}\pi\right) Z \cos(\theta) +$$

$$+ Y \sin\left(\theta - \frac{28}{180}\pi\right) + Z \sin\left(\theta - \frac{62}{180}\pi\right)$$

Il sistema in Eq. (3) è costituito da equazioni differenziali lineari a coefficienti costanti: pertanto esso ammette una soluzione in forma chiusa. Questa può essere determinata in maniera completa, nei termini degli spostamenti angolare, radiale e tangenziale ϕ , η ed u , in funzione della coordinata polare θ , imponendo le condizioni al contorno. Queste sono relative alla congruenza della struttura isostatica equivalente rispetto ai vincoli esterni. Nella fattispecie, occorre imporre che lo spostamento tangenziale sia nullo agli estremi e così pure che sia pari a zero la rotazione per $\theta = \pi/2$, oltre alle condizioni di continuità su tutti gli spostamenti internamente alla struttura. Un'apposita routine in Matlab (Versione 7.9.0.529 (R2009b)) è stata sviluppata per presiedere all'integrazione delle equazioni differenziali (Eq. (3), tenendo conto delle espressioni nelle Eqq. (4-9)), all'imposizione delle condizioni al contorno con relativa determinazione delle costanti di integrazione e per la scrittura delle equazioni di congruenza (Eq. (2)). Queste ultime vengono quindi risolte per la valutazione delle incognite iperstatiche, mentre semplici considerazioni di equilibrio conducono alla stima delle altre reazioni vincolari H_B , V_A e T_A .

Valutazione delle reazioni sull'anello in direzione verticale

Le forze in gioco, dovute al peso proprio dell'anello ed alle masse concentrate agiscono ortogonalmente rispetto allo sviluppo circonferenziale dell'anello stesso. Nella Fig. (6) viene pertanto mostrato lo sviluppo in piano per tutta l'estensione dell'anello. Anche in questo caso, come già discusso al paragrafo precedente la simmetria geometrica e di carico permette di passare gradualmente a configurazioni semplificate, sezionando l'anello lungo i piani di simmetria. In ultima analisi, può essere considerato anche in questo caso solo un quarto di anello.

Tale struttura è stata studiata per la determinazione del sistema reattivo. Essa risulta due volte iperstatica: per la sua risoluzione si è adottato il "metodo delle forze" [7], assumendo come incognite iperstatiche i due momenti S e W . Le relative equazioni di congruenza sono mostrate in Eq. (10): anche in questo per esplicitarle, occorre ricavare l'equazione della linea elastica. Questa può essere valutata risolvendo il sistema di equazioni differenziali illustrato in Eq. (11), applicando le condizioni al contorno di continuità sugli spostamenti fra i tre tronchi di trave e di abbassamento nullo in corrispondenza degli appoggi. Si indicano rispettivamente con ϕ la rotazione e con v lo spostamento verticale. Per comodità, le equazioni differenziali in Eq. (11) sono espresse nella coordinata polare θ , che individua la posizione angolare lungo la porzione di anello.

$$\begin{cases} \phi(\theta=0)=0 \\ \phi\left(\theta=\frac{\pi}{2}\right)=0 \end{cases} \quad (10)$$

$$\begin{cases} \frac{M_x(\theta)R_G}{EI_x} = -\frac{d}{d\theta}\phi(\theta) \\ \phi(\theta)R_G = \frac{d}{d\theta}v(\theta) \end{cases} \quad (11)$$

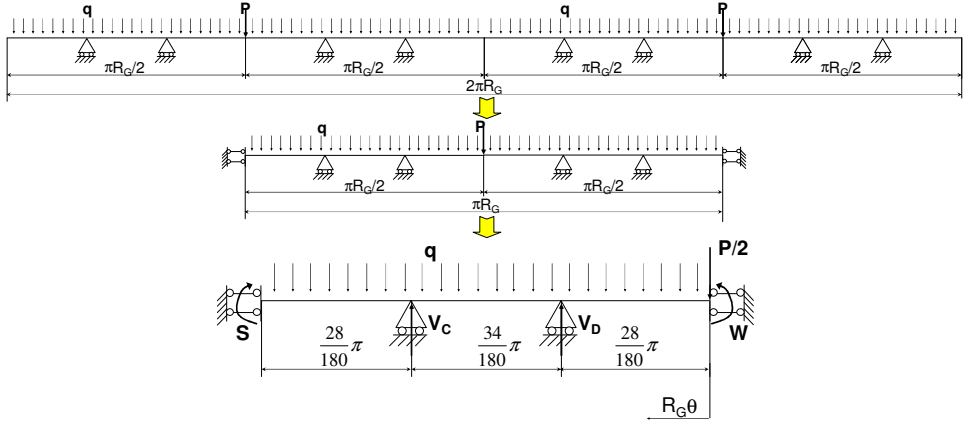


Figura 6. Sviluppo in piano del prospetto dell'anello e determinazione per simmetria della struttura iperstatica effettivamente risolta.

In Eq. (11) il termine I_x indica il momento inerziale di area della sezione, valutato rispetto ad un asse orizzontale baricentrico. Con $M_x(\theta)$ si indica il momento flettente dovuto alle forze verticali (assunto positivo quando tende le fibre inferiori): esso risulta così esprimibile, Eqq. (12-14):

$$M_x(\theta) = W - \frac{P}{2} R_G \theta - \frac{q R_G^2 \theta^2}{2} \quad \text{per } 0 < \theta \leq \frac{28}{180} \pi \quad (12)$$

$$\begin{aligned} M_x(\theta) = & W - \frac{W}{\frac{34}{180} \pi} \left(\theta - \frac{28}{180} \pi \right) + \\ & + \frac{S}{\frac{34}{180} \pi} \left(\theta - \frac{28}{180} \pi \right) - \frac{P}{2} R_G \theta + \\ & + \frac{P}{2} \frac{31}{17} R_G \left(\theta - \frac{28}{180} \pi \right) - \frac{q R_G^2 \theta^2}{2} + \\ & + q R_G^2 \frac{\pi}{4} \left(\theta - \frac{28}{180} \pi \right) \end{aligned} \quad \text{per } \frac{28}{180} \pi < \theta \leq \frac{62}{180} \pi \quad (13)$$

$$\begin{aligned} M_x(\theta) = & S - \frac{q R_G^2 \theta^2}{2} + q R_G^2 \frac{\pi}{4} \left(\theta - \frac{28}{180} \pi \right) + \\ & + q R_G^2 \frac{\pi}{4} \left(\theta - \frac{62}{180} \pi \right) \end{aligned} \quad \text{per } \frac{62}{180} \pi < \theta \leq \frac{\pi}{2} \quad (14)$$

Anche in questo caso è stata sviluppata espressamente una routine in Matlab, che permette di integrare le equazioni differenziali di Eq. (11), tenendo conto delle espressioni in Eqq. (12-14), di imporre le condizioni al contorno e di determinare le costanti di

integrazione. Segue quindi la scrittura e risoluzione del sistema delle equazioni di congruenza (Eq. (10)) e la determinazione dei momenti incogniti W ed S . Ulteriori considerazioni di equilibrio portano infine al calcolo delle reazioni sui carrelli V_C e V_D .

Valutazione delle caratteristiche della sollecitazione sull'anello e determinazione dello stato tensionale e deformativo

Una volta risolte le strutture iperstatiche, tramite gli approcci analitici mostrati ai paragrafi precedenti, è possibile valutare le caratteristiche della sollecitazione sul quarto di anello. Sulla base infatti di tutte le componenti reattive precedentemente calcolate, è possibile esprimere gli andamenti dei momenti flettenti lungo le direzioni x e y (M_x ed M_y) e dello sforzo normale (N), anche sulla scorta delle rispettive espressioni analitiche, contenute rispettivamente nelle Eqq. (12-14), Eqq. (4-6) ed Eqq. (7-9). La presente trattazione prescinde dalle componenti di taglio, che danno un contributo trascurabile alla sollecitazione della struttura: i rispettivi andamenti non sono quindi presi in considerazione. Un'altra semplificazione ha portato a trascurare la leggera torsione dell'anello in corrispondenza dell'applicazione dei carichi concentrati. Tale momento torcente è dovuto al disassamento fra il baricentro delle masse e la linea d'asse tangenziale dell'anello. Il valore contenuto di questo disassamento, molto minore rispetto al diametro dell'anello, ha autorizzato a trascurare questa ulteriore componente di sollecitazione.

La conoscenza delle tre caratteristiche di momento flettente nelle direzioni orizzontale e verticale e di sforzo normale ha quindi permesso di impostare il calcolo della tensione che si sviluppa lungo la linea d'asse tangenziale dell'anello stesso. Per il calcolo di questa, si è applicata l'Eq. (15) con riferimento alla Fig. (7), che raccoglie le informazioni sulle caratteristiche geometriche dell'anello e sulle sollecitazioni ivi presenti. Si può osservare che si tratta del classico modello a tre termini per la flessione deviata composta, che è stato però corretto, per tenere conto del fatto che il momento flettente lungo l'asse y fa inflettere l'anello come trave ad asse curvilineo.

$$\sigma = \frac{N}{A} + \frac{M_x}{I_x} y - \frac{M_y}{eA} \frac{x+e}{R_N + x+e} \quad (15)$$

In tale relazione i termini A ed I_x assumono il significato già noto, mentre R_N indica il raggio che individua la posizione dell'asse neutro. Il termine e si riferisce alla cosiddetta eccentricità, ossia alla differenza fra il raggio baricentrico R_G di Fig. (5) ed il raggio R_N (si veda anche la Fig. (7)). Tale raggio, tenendo conto della trattazione classica per le travi ad asse curvo, può essere così calcolato (si trascura la lavorazione in corrispondenza della guida inferiore):

$$R_N = \frac{A}{\int_A \frac{dA}{r}} = \frac{A}{L_{est.} \cdot \lg\left(\frac{R_{est.}}{R_{int.}}\right) - L_{int.} \cdot \lg\left(\frac{R_{est.} - \frac{L_{est.} - L_{int.}}{2}}{R_{int.} + \frac{L_{est.} - L_{int.}}{2}}\right) - h \cdot \lg\left(\frac{R_{est.}}{R_{est.} - s}\right)} \quad (16)$$

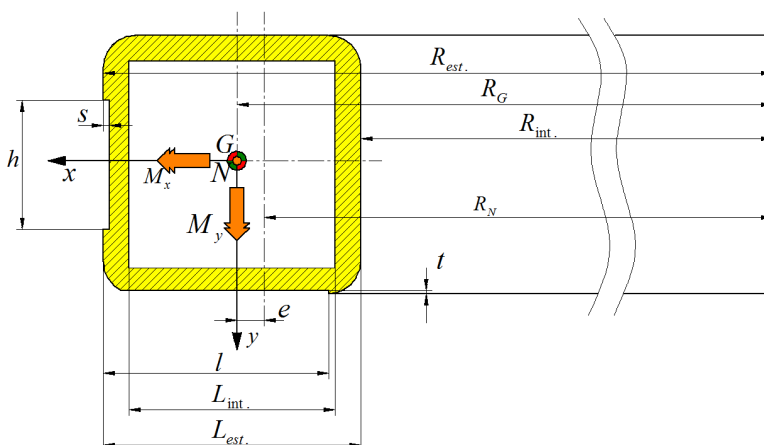


Figura 7. Sezione dell'anello, con indicazione delle quote caratteristiche e delle sollecitazioni.

Nell'Eq. (16) dA indica il differenziale della sezione A , mentre r rappresenta la coordinata radiale generica. Gli altri simboli hanno il significato chiarito nella Fig. (7). La trattazione è completata dalla determinazione dell'equazione della linea neutra, che in tal caso risulta essere una curva con andamento comunque molto prossimo a quello lineare. Tale equazione è ottenibile eguagliando a zero l'espressione di Eq. (15) ed è esplicitata in Eq. (17).

$$y = \frac{I_x}{M_x} \left(-\frac{N}{A} + \frac{M_y}{eA} \frac{x+e}{R_N + x+e} \right) \quad (17)$$

4. RISULTATI DEL MODELLO ANALITICO

Il modello descritto nel Par. 3 è stato più volte applicato per determinare le caratteristiche della sollecitazione e, successivamente, gli stati di tensione e deformazione nella sezione dell'anello. Nelle immagini nella Fig. (8) sono mostrati gli andamenti dei momenti flettenti e dello sforzo normale, in funzione della coordinata angolare θ . E' da notare che il momento lungo l'asse y M_y e lo sforzo normale N dipendono dalle componenti centrifughe e di conseguenza dalla velocità angolare dell'anello. Il modello analitico è stato quindi utilizzato in cascata, considerando diverse velocità angolari, le stesse cui sono state eseguite le prove sperimentali, che saranno descritte nei Parr. 6 e 7. I risultati mostrati si riferiscono alla velocità angolare massima, pari a 53 rpm.

Il calcolo delle tensioni è stato preceduto dalla valutazione dell'andamento dell'asse neutro, in accordo con l'Eq. (17). Stante la configurazione delle Figg. (5-6), si sono intesi come punti notevoli quelli in corrispondenza dell'applicazione dei carichi concentrati ($\theta = 0^\circ$), degli appoggi ($\theta = 28^\circ$ e $\theta = 62^\circ$) ed infine il punto in quadratura rispetto al primo, anch'esso al centro della campata più lunga ($\theta = 90^\circ$).

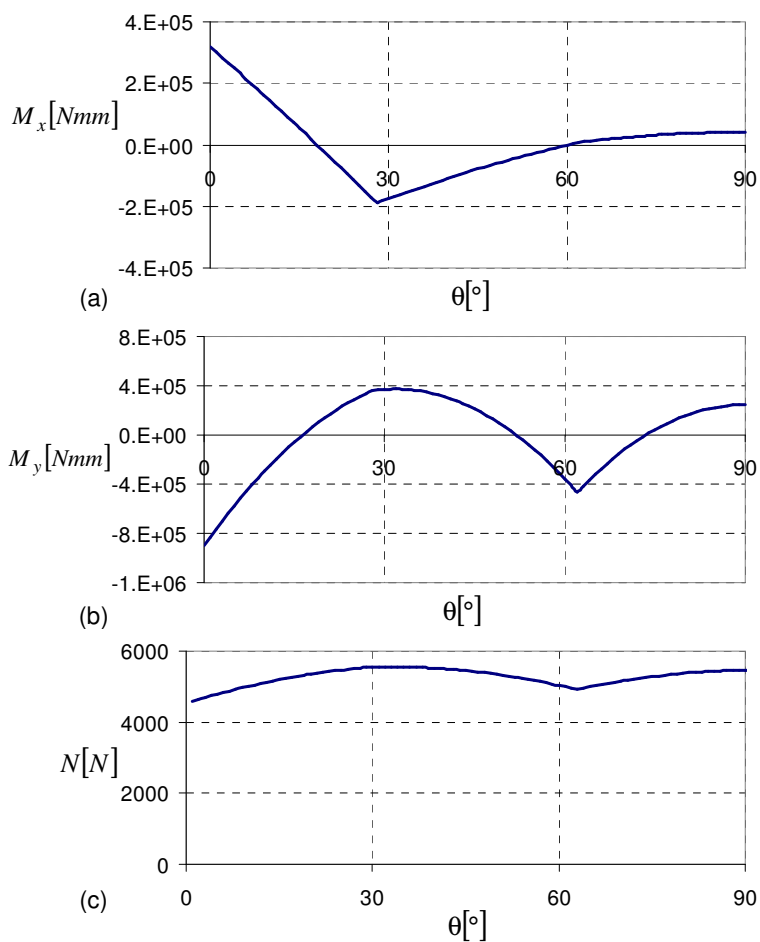


Figura 8. Andamento di M_x , M_y e N in funzione di θ .

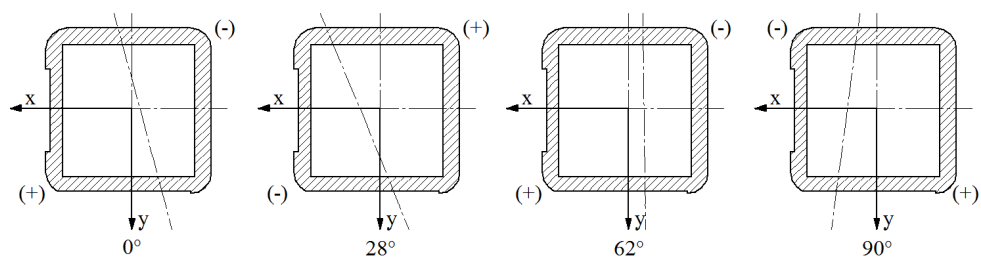


Figura 9. Disposizione dell'asse neutro e zone in trazione ed in compressione a diversi valori di θ .

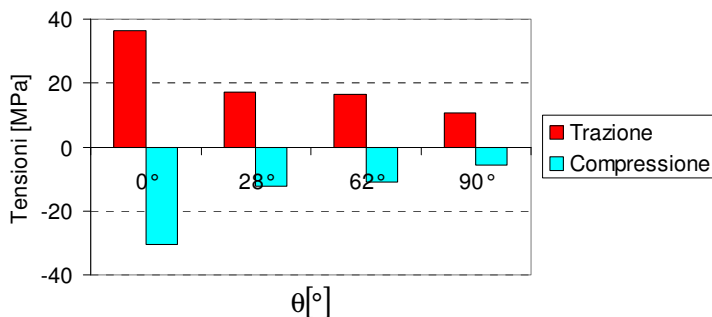


Figura 10. Valori massimi di tensione in trazione e compressione per diversi valori di θ .

La Fig. (9) mostra gli andamenti delle linee neutre per le sezioni appena evidenziate. I simboli (+) e (-) indicano i punti maggiormente sollecitati a trazione e compressione rispettivamente. I valori quantitativi corrispondenti sono infine mostrati nella Fig. (10).

5. DISCUSSIONE E PREPARAZIONE ALLE PROVE SPERIMENTALI

Sulla base dei risultati ottenuti dal modello analitico, sia dal punto di vista qualitativo della distribuzione degli sforzi, che da quello quantitativo dell'intensità degli stessi, è stato possibile trarre una serie di considerazioni, che hanno rappresentato il preludio all'attività sperimentale. Una volta stabilite le posizioni di applicazione degli estensimetri, il modello analitico è stato ulteriormente migliorato, per cercare di prevedere i risultati delle prove sperimentali.

Discussione dei risultati del modello analitico

L'analisi dei risultati nelle Figg. (8-10) porta ad una serie di osservazioni che sono così riassumibili.

- La Fig. (9), relativa agli andamenti degli assi neutri, mostra come questi abbiano un'inclinazione quasi verticale, o comunque più vicina alla verticalità che non all'orizzontalità. Questo indica che la componente prevalente è quella dovuta ai carichi centrifughi.
- L'analisi della Fig. (8 a-b), oltre che della Fig. (10), mostra come le sollecitazioni, i momenti flettenti e le tensioni normali che ne derivano, siano nettamente superiori alla posizione angolare $\theta = 0^\circ$, ossia in corrispondenza dei carichi concentrati radiali e verticali. La tensione nel punto in quadratura ($\theta = 90^\circ$), pure esso al centro della campata più lunga, ma soggetto solo ai carichi distribuiti, anch'essi radiali e verticali, è molto minore. Considerando anche l'osservazione del punto precedente, si è portati a ritenere che, delle quattro tipologie di carico illustrate al Par. 2, il carico concentrato centrifugo rappresenti la componente più gravosa.

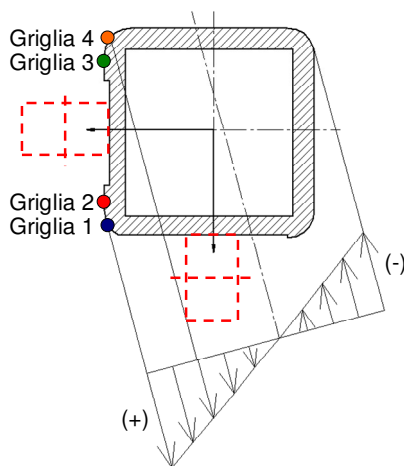


Figura 11. Andamento qualitativo di tensioni e deformazioni a $\theta = 0^\circ$ e posizioni scelte per l'applicazione degli estensimetri elettrici.

- Dal punto di vista quantitativo il picco di tensione stimato è intorno ai 36 MPa: tale valore corrisponde ad una deformazione stimata in circa $170 \mu\epsilon$: si tratta quindi di un valore che è rilevabile anche per via sperimentale, tramite la conduzione di prove estensimetriche.

L'indicazione che deriva quindi dai risultati del modello analitico è quella di disporre estensimetri elettrici in corrispondenza della sezione dove agiscono i carichi concentrati ($\theta = 0^\circ$), o comunque, ove possibile, in prossimità di questa. La Fig. (11) mostra, a livello qualitativo, la distribuzione delle tensioni e delle deformazioni in corrispondenza di tale sezione e si nota che tutta la parte esterna risulta in trazione con il picco di tensione in corrispondenza del vertice inferiore. Sulla base di questa osservazione, si sono individuate quattro posizioni per l'incollaggio di estensimetri a resistenza, prevedendo l'impiego di sensori monogriglia, disposti in direzione tangenziale. Dalle posizioni possibili si sono dovute stralciare quelle in corrispondenza dei raccordi, dove si ha una doppia curvatura, e quelle in corrispondenza delle guide laterale ed inferiore, dove agiscono le ruote gommate.

Modello agli elementi finiti

Per avere un'ulteriore conferma sull'efficacia dell'applicazione di estensimetri elettrici nelle posizioni indicate, e per meglio rendersi conto del gradiente di tensione al passare dal bordo esterno inferiore a quello superiore dell'anello, è stato messo a punto un modello agli elementi finiti. Va anche sottolineato che il modello analitico trascurava alcuni aspetti, quali il momento torcente in corrispondenza dei carichi concentrati e gli sforzi di taglio. Il modello agli elementi finiti è stato realizzato tramite il software Ansys (Versione 12.0), a partire da un modello solido costruito tramite il software SolidEdge (Versione 18.00.00.69). Su di esso sono state applicate le quattro tipologie di carico, discusse al Par. 2, tenendo conto dell'effettiva posizione dei baricentri delle masse concentrate. Si è proceduto quindi al vincolamento in otto punti, con vincoli in direzione radiale e verticale. La struttura è stata quindi discretizzata, tramite elementi tetraedrici, per un totale di 76.348 elementi e 151.540 nodi.

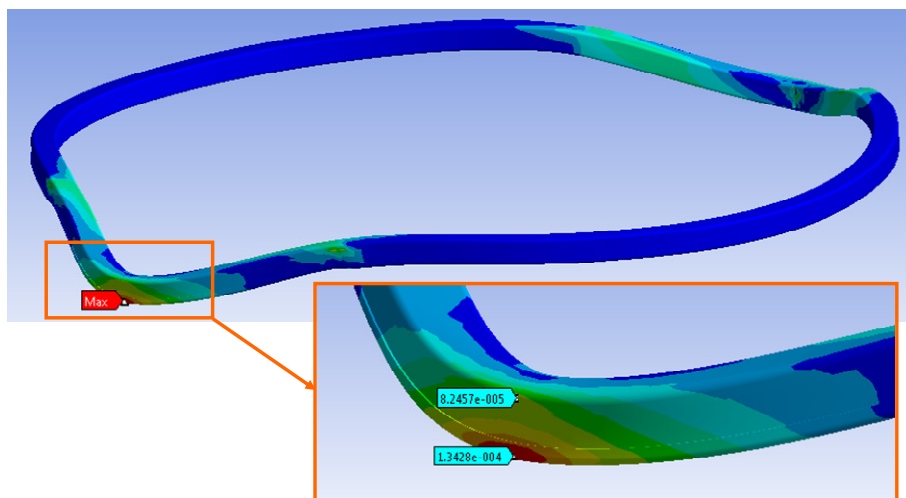


Figura 12. Distribuzione delle deformazioni principali nell'anello secondo il modello FEM alla massima velocità di rotazione.

Il risultato della simulazione (operata per la massima velocità di rotazione) è mostrato in Fig. (12), con riferimento alle deformazioni principali massime ed un dettaglio sulle locazioni degli estensimetri elettrici chiamati a misurare tali deformazioni. La distribuzione risulta coerente a quella mostrata in Fig. (11), con un valore di picco al bordo inferiore di circa $130 \mu\epsilon$, mentre al bordo superiore si ha una deformazione di circa $80 \mu\epsilon$, comunque ben misurabile con estensimetri convenzionali. La porzione di anello in quadratura rispetto a tale sezione risulta invece praticamente scarica, coerentemente con quanto prima osservato. Si può ravvisare che il modello analitico precedentemente descritto si presenta più conservativo, portando a valori leggermente maggiori rispetto a quelli che scaturiscono dalla simulazione agli elementi finiti. Questo è tra l'altro dovuto al fatto che i carichi concentrati sono nella realtà e nel modello numerico leggermente distribuiti attorno al punto di applicazione nominale.

Aggiornamento del modello analitico per poter prevedere le letture estensimetriche

Una volta stabilite le posizioni di applicazione degli estensimetri e previsto, come verrà esposto nel dettaglio al Par. 6, di effettuare le misurazioni durante più giri completi dell'anello, si è cercato di sviluppare, sulla scorta di quello esposto al Par. 3, una nuova versione del modello analitico. Visto che l'obiettivo finale era quello di validare il modello, tramite confronto con i risultati sperimentali, è sembrato opportuno modificarlo, in modo tale che potesse prevedere a calcolo le deformazioni lette da ciascun estensimetro durante una rotazione completa della struttura. In tale modo sarebbe stato molto più immediato procedere al confronto qualitativo e quantitativo delle grandezze rilevate e stimarne lo scostamento percentuale. L'idea di fondo è stata quella di iterare il procedimento descritto nel Par. 3 per numerosi passi angolari, corrispondenti allo spostamento γ durante la rotazione, valutando ad ogni passo le caratteristiche della sollecitazione della struttura e tensioni e deformazioni in corrispondenza della sezione di applicazione degli estensimetri.

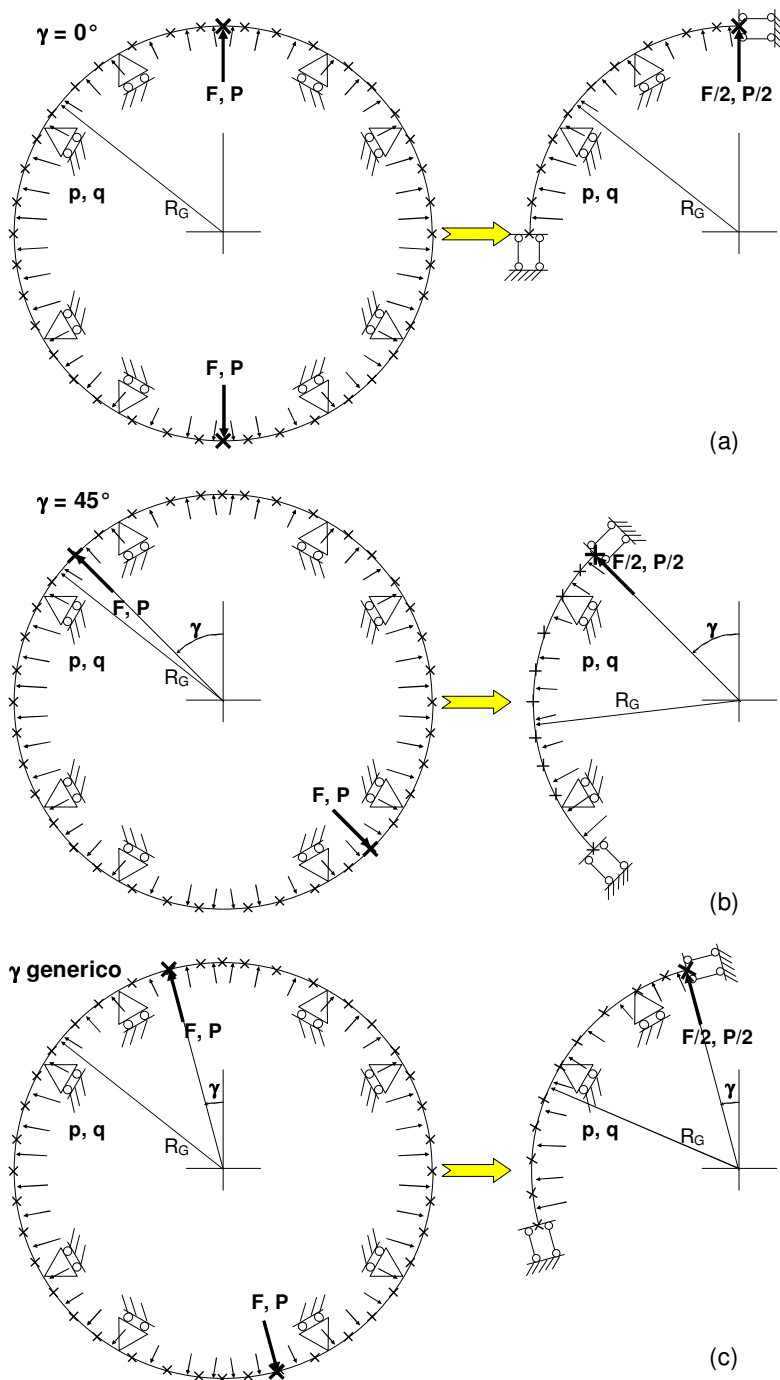


Figura 13. Determinazione delle strutture equivalenti nel modello analitico aggiornato per diversi valori dell'angolo di rotazione γ .

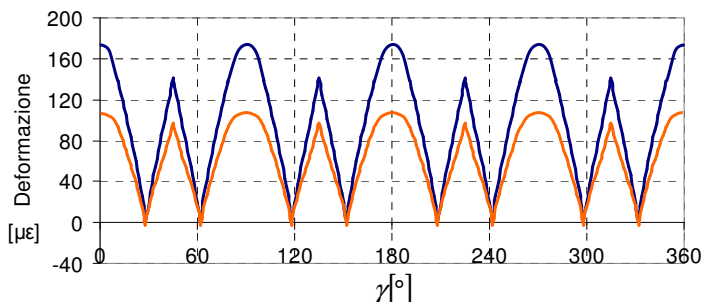


Figura 14. Previsione sugli andamenti delle deformazioni misurate dagli estensimetri in posizioni estreme secondo il modello analitico aggiornato.

La Fig. (13) mostra come ad ogni passo angolare venga considerata una nuova struttura equivalente, corrispondente ad un quarto di anello, che viene studiata, ripercorrendo quanto esposto nel Par. 3. Guardando la stessa figura, si nota che il passaggio dalla struttura completa chiusa a quella semplificata è rigoroso per $\gamma = 0^\circ$ (carichi concentrati verticale e radiale al centro della campata massima, Fig. (13 a), si tratta del caso già esaminato nel Par. 3), per $\gamma = 45^\circ$ (carichi concentrati al centro della campata più piccola, fra due appoggi, Fig. (13 b)), e così via con una periodicità di 45° . In corrispondenza di queste posizioni angolari considerazioni di simmetria permettono di semplificare la struttura: si tratta delle posizioni più significative, in quanto è in esse che si registrano i massimi relativi ed assoluti per le tensioni e deformazioni sulla sezione strumentata. In tutte le altre posizioni angolari, per γ generici (Fig. (13 c)), non si ha una perfetta simmetria, il che a rigore non permetterebbe la semplificazione della struttura, complicando fortemente la trattazione analitica. In ragione del fatto che i valori di picco sono correttamente individuati e che al passaggio dei carichi concentrati in corrispondenza degli appoggi la sollecitazione è ragionevolmente molto bassa, è parso ragionevole, in corrispondenza di tali posizioni, ricavare la struttura equivalente in maniera semplificata. Nella fattispecie si è supposto che in corrispondenza dei carichi concentrati la rotazione (angoli φ e ϕ secondo la notazione precedentemente introdotta) sia trascurabile: di conseguenza in questi punti si è posto un vincolo a doppio pendolo, che impedisce le rotazioni. Altri casi particolari si hanno quando i carichi concentrati si trovano ad insistere proprio in corrispondenza degli appoggi (ad esempio per $\gamma = 28^\circ$ e $\gamma = 62^\circ$): secondo la schematizzazione scelta, in tali punti vengono poste combinazioni di appoggi e doppi pendoli, ossia degli incastri, su cui si scaricano direttamente i carichi concentrati. La schematizzazione appare coerente, dato che nella realtà i carichi si scaricano effettivamente sulle ruote gommate, che impediscono ogni spostamento in direzione verticale e trasversale, mentre eventuali rotazioni potrebbero essere dovute all'influsso dei carichi distribuiti. Essendo però tali carichi di entità piuttosto limitata, come già ravvisato precedentemente, le rotazioni saranno verosimilmente praticamente nulle.

La Fig. (14) mostra l'andamento previsto per le deformazioni rilevate dagli estensimetri nelle posizioni estreme (numeri delle griglie 1 e 4), quello sul bordo esterno inferiore, che fornisce la lettura maggiore e quello al bordo superiore. L'immagine si riferisce alla velocità angolare più elevata della macchina, 53 rpm, ma ulteriori elaborazioni sono state effettuate anche per velocità minori. Osservando l'andamento da un punto di

vista qualitativo, si notano anche dei tratti distintivi, che saranno ravvisati anche dall'esame dei dati sperimentali. Si riscontrano dei picchi in corrispondenza del centro della campata più lunga, successivamente si hanno due minimi vicini al transito sugli appoggi, in cui tensione e deformazione sono quasi nulle, e si ha infine un piccolo picco a separare questi due minimi, in corrispondenza del centro della campata minore.

6. PROCEDURA SPERIMENTALE

Le prove sperimentali sul campo sono state effettuate sulla macchina rappresentata nelle Figg. (2-3), dopo la strumentazione dell'anello rotante. Tale strumentazione è avvenuta, secondo quanto riportato nel Par. 5, in corrispondenza della sezione dove sono posti i carichi concentrati, ossia dove si trovano il carrello porta-bobina ed il gruppo pre-stiro. Nelle fasi preparatorie alle prove sono stati applicati (Fig. (15 a)) quattro estensimetri elettrici monogriglia (CEA-XX-125-UW-120, Produttore: Vishay, Malvern, PA, Stati Uniti) nelle posizioni indicate in Fig. (11) con orientazione longitudinale lungo la direzione tangenziale.

Per quanto riguarda il sistema di acquisizione, la principale difficoltà era data dalla necessità di acquisire le letture estensimetriche da un organo in movimento, rotante ad una velocità non trascurabile. In questi casi esistono dispositivi a contatto strisciante per la trasmissione del segnale, che però, se non adeguatamente schermati, possono comportare l'ingresso di rumore sul segnale acquisito. In più, questi strumenti sono solitamente poco flessibili: sviluppati essenzialmente per le misure di coppia su alberi rotanti (realizzazione di torsimetri), sono poco adattabili ad applicazioni diverse, come quella in studio, a meno di non costruire appositamente attrezzature aggiuntive non commerciali. Nel presente caso si è preferito ricorrere ad una centralina per la trasmissione wireless (NI WLS-9237 IEEE 802.11b/g, Produttore: National Instruments, Austin, TX, Stati Uniti). Tale centralina è stata bloccata con l'ausilio di fascette in corrispondenza di un montante solidale con l'anello rotante (Fig. (15 c)). A questa sono stati collegati 4 cavi, uno per canale, ossia per estensimetro, a loro volta connessi a piattaforme NI 9945 (Produttore: National Instruments), specifiche per la realizzazione di circuiti a quarto di ponte di Wheatstone (Fig. (15 b)).

A questo punto, un'ulteriore difficoltà è stata data dalla necessità di collegare i terminali elettrici degli estensimetri con tali piattaforme, che erano fissate sul lato interno dell'anello, in prossimità della sua superficie superiore. Infatti, osservando la Fig. (11), si può comprendere che, mentre i sensori posti superiormente potevano essere agevolmente collegati, i fili di collegamento degli altri due sarebbero dovuti forzatamente transitare per una delle due guide dell'anello, rischiando di interferire con il passaggio sulle ruote di appoggio. Per minimizzare questi rischi, si sono utilizzati fili in rame pre-isolati, con spessore particolarmente ridotto (dell'ordine del decimo di mm), che attraversano la guida inferiore e sono qui bloccati con opportuno adesivo (Fig. (15 a)). Sono quindi saldati ad una basetta intermedia, da cui si dipartono ordinari cavi a quattro poli fino alle piattaforme. L'impiego di tali fili ha rappresentato un buon compromesso fra la necessità di un collegamento sicuro ed affidabile dal punto di vista della qualità del segnale acquisito e quella di non interferire sul regolare funzionamento della macchina. Visto il loro basso spessore, essi possono infatti essere calpestati dalle ruote senza particolari inconvenienti né a livello funzionale, né a livello di accuratezza sulla misura.

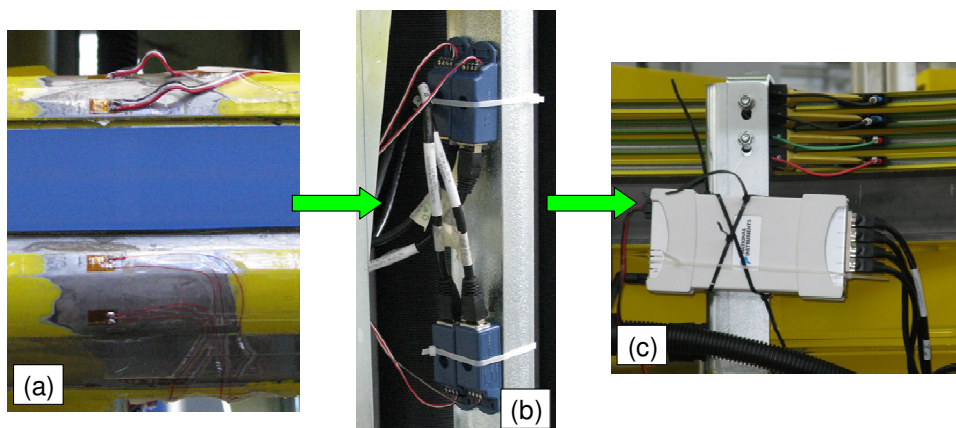


Figura 15. Catena di misura per le prove sperimentali.

La Fig. (15) mostra in sequenza i componenti della catena di misura appena descritti. La centralina si collega in rete con il computer portatile, deputato all'acquisizione e memorizzazione dei dati. Per coordinare le operazioni sperimentali, si è costruito espressamente un programma LabView, tramite il quale si possono comandare l'inizio dell'acquisizione e il salvataggio dei dati, monitorando in tempo reale i segnali elettrici acquisiti, automaticamente convertiti in segnali di deformazione. Le prove sono state effettuate alle velocità di 53 rpm, 43, 33, 23, 13, 0,5, facendo compiere diversi giri all'anello. Per ragioni di rilevanza statistica dei risultati, ogni prova, a ciascun livello di velocità, è stata ripetuta per tre volte.

7. RISULTATI SPERIMENTALI

In ognuna delle prove effettuate sono state misurate le deformazioni in corrispondenza degli estensimetri, campionando i segnali alla frequenza di 50 Hz, sufficiente per una buona descrizione dell'andamento delle grandezze indagate nel corso della rotazione. Successivamente, tutti i risultati sono stati elaborati per la stima del livello di tensione nella sezione considerata: allo scopo si è utilizzata la legge di Hooke, supponendo il valore di 210 GPa per il modulo elastico, già utilizzato anche nella redazione del modello analitico (Par. 3).

La Fig. (16) si riferisce ad un'acquisizione a 53 rpm, i colori si riferiscono ai quattro canali di misura, ognuno dei quali è relativo alle griglie estensimetriche schematizzate in Fig. (11). Si nota come in essa siano riscontrabili gli stessi caratteri distintivi, già evidenziati nel Par. 5, a proposito dei risultati analitici. Come evidenziato nella stessa figura, si hanno i picchi più alti in corrispondenza del transito per il centro della campata più lunga, coppie di minimi al passaggio sugli appoggi e picchi di minore entità nella breve campata fra le coppie di appoggi. Si rimarca inoltre il fatto che l'andamento dei segnali acquisiti appare regolare: non compaiono cioè discontinuità, che possano essere imputabili a disturbi sulla catena di misura. Inoltre, in nessuna prova si sono osservati effetti dinamici nei transitori all'avviamento o alla fermata.

I diagrammi in Fig. (17) riportano le elaborazioni in termini di tensioni alle due velocità più alte (53 rpm, Fig. (17 a) e 43 rpm, Fig. (17 b)).

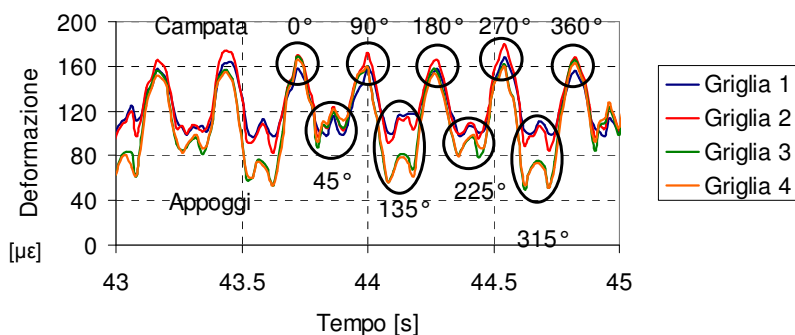


Figura 16. Andamento delle deformazioni durante la rotazione dell'anello a 53 rpm.

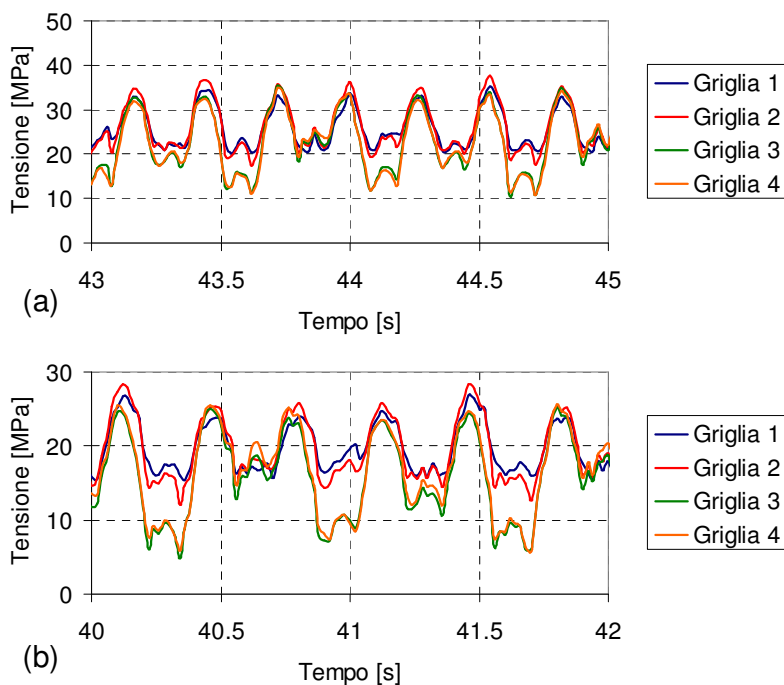


Figura 17. Andamento delle tensioni durante la rotazione dell'anello a 53 rpm (a) ed a 43 rpm (b).

8. VALIDAZIONE DEL MODELLO ANALITICO

Nel presente lavoro sono stati presentati due modelli analitici: il primo permette di valutare lo stato tensionale e deformativo dell'anello, quando questo si trova nella posizione più critica (carichi concentrati al centro delle campate più lunghe), il secondo, evoluzione del primo, si focalizza sullo stato tensionale e deformativo in corrispondenza dei carichi concentrati e monitora la sua evoluzione durante un giro completo dell'anello. Entrambi i

modelli sono tra loro legati, i risultati del secondo per $\gamma = 0^\circ$ (ad inizio rotazione) sono gli stessi del primo per $\theta = 0^\circ$. Pertanto, si è validato il secondo, intendendo con questo validato conseguentemente anche il primo. Il confronto è stato operato fra gli andamenti, sperimentale e simulato, delle tensioni e deformazioni rilevate dagli estensimetri. Successivamente, si sono comparati anche i valori di picco (quelli al centro della campata) tramite un diagramma di verità.

L'acquisizione è stata effettuata solo sui quattro canali di sollecitazione e non si è potuto prevedere un ulteriore segnale per monitorare la posizione angolare dell'anello rotante. I diagrammi sperimentali e simulati, hanno comunque potuto essere fasati e sovrapposti, portando a coincidere gli elementi distintivi, in particolare i picchi massimi, prima evidenziati. Si mostrano in Fig. (18) i confronti degli andamenti delle deformazioni alle due velocità maggiori, 53 (Fig. (18 a)) e 43 (Fig. (18 b)) rpm. I diagrammi si riferiscono al punto strumentato, in cui su registra la massima sollecitazione nella sezione (griglia 1). Si nota una buona corrispondenza, soprattutto per quanto riguarda i picchi massimi al centro della campata, oltre che sull'andamento globale.

Il confronto fra i picchi massimi è approfondito nel diagramma di verità in Fig. (19), che pone a confronto diretto la massima deformazione rilevata e la massima stimata per via analitica. Come si può vedere, le differenze sono minime con un errore mediamente di pochi punti percentuali.

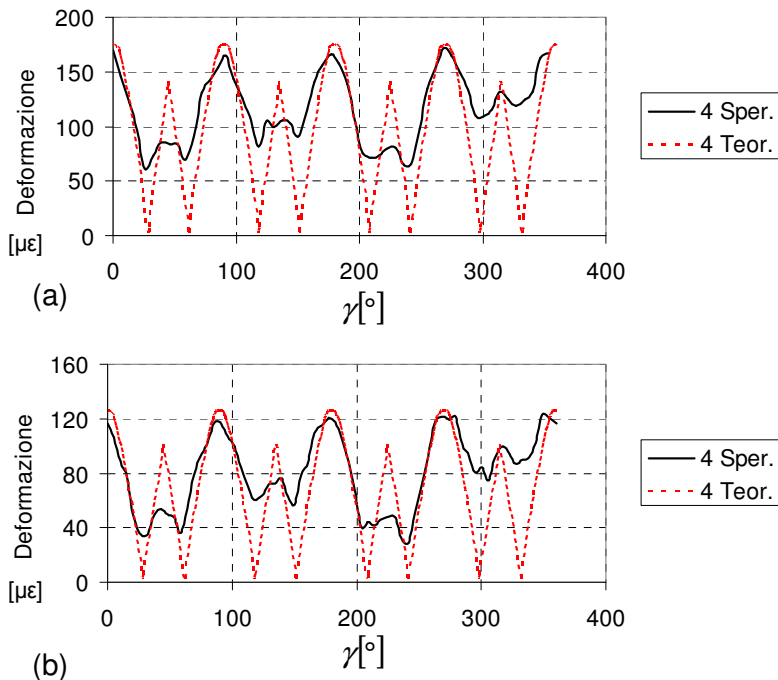


Figura 18. Confronto fra le deformazioni nel punto più sollecitato ricavate per via analitica e rilevate sperimentalmente.

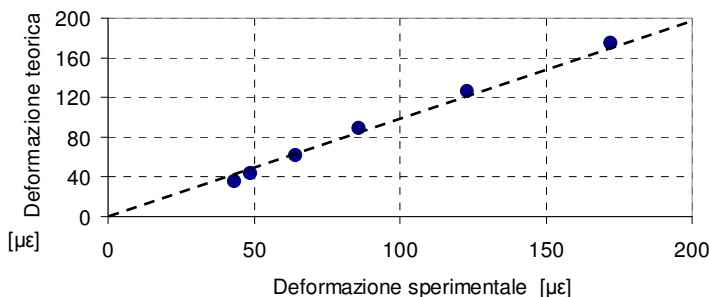


Figura 19. Diagramma di verità per porre a confronto le deformazioni di picco, valutate per via analitica e sperimentale.

9. CONCLUSIONI

A seguito degli studi, teorici e sperimentali ivi descritti, si può rimarcare quanto segue.

- L'analisi dello stato dell'arte ha mostrato come le macchine cellofanatrici possano assumere diverse configurazioni (ad asse orizzontale o verticale, a due o quattro montanti), ma che la caratteristica comune resta legata alla presenza dell'anello rotante, che realizza materialmente l'imballaggio. Nonostante l'importanza di questo organo, non si riscontrano in letteratura studi teorico-analitici, che mirino alla sua modellazione, tenendo conto di tutti i carichi e vincoli presenti.
- Concentrandosi su una macchina ad elevata produttività, che, per conferire maggiore rigidità all'anello, utilizza una struttura fortemente iperstatica, si è sviluppato un modello analitico, che ha condotto alla valutazione delle reazioni iperstatiche, a quella delle caratteristiche della sollecitazione ed alla stima dei livelli di tensione e deformazione nella sezione più sollecitata.
- I risultati analitici hanno mostrato come le sollecitazioni centrifughe, soprattutto se concentrate, abbiano un impatto maggiore rispetto a quelle gravitazionali. Questo ha permesso di individuare le posizioni in cui applicare sensori di deformazione per una validazione del modello sul campo. L'entità della deformazione stimata, intorno ai 170 µε alla massima velocità, ha indotto verso la scelta della tecnica estensimetrica.
- Lo stesso modello è stato ulteriormente perfezionato, per renderlo in grado di prevedere a calcolo le letture estensimetriche nell'arco di una rotazione completa dell'anello.
- Le prove sperimentali sono state eseguite, acquisendo le deformazioni di quattro estensimetri elettrici posti sulla sezione più critica dell'anello. Le principali difficoltà sono state legate alla trasmissione wireless del segnale fra le postazioni rotorica (anello in moto) e statorica (calcolatore per l'elaborazione dei dati), ed alla realizzazione del cablaggio degli estensimetri, per cui si è fatto ricorso ad una soluzione innovativa ed originale.
- I risultati delle prove, rapportati a quelli analitici, hanno evidenziato un ottimo accordo, sia dal punto di vista qualitativo che quantitativo, con errori di pochi punti percentuali (mediamente intorno al 4%).

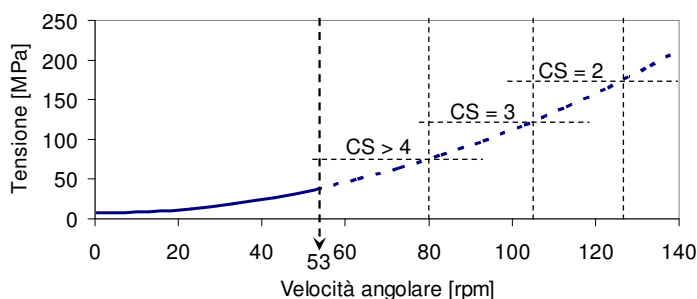


Figura 20. Legame fra tensione massima nell'anello e velocità di rotazione.

- Il modello validato potrà essere impiegato a supporto dell'attività progettuale e per eventualmente ridefinire i parametri di funzionamento, con particolare riferimento alla velocità angolare di rotazione. La Fig. (20) mostra come alla velocità di 53 rpm le tensioni massime siano ancora molto limitate rispetto al limite allo snervamento dell'anello (355 MPa, [8]). Di conseguenza, la velocità può essere ragionevolmente aumentata, almeno fino a valori intorno ai 70-80 rpm, già raggiunti da macchine di caratteristiche simili [9], mantenendo un coefficiente di sicurezza (CS) molto elevato. Ulteriori incrementi con CS sufficientemente alti potrebbero essere possibili, anche se dovranno essere attentamente indagati, per evitare fenomeni di risonanza previsti alle velocità particolarmente alte.

Ringraziamenti

Gli autori ringraziano l'Ing. Paolo Proli della Seconda Facoltà di Ingegneria dell'Università degli Studi di Bologna ed il personale della Ditta Mecwrap (<http://www.mecwrap.com>) per il supporto nell'esecuzione delle prove sperimentali.

BIBLIOGRAFIA

- [1] Cowley, P. R., 1985. "The Experience Curve and History of the Cellophane Business". *Long Range Planning*, **18**(6), pp. 84–90.
- [2] Mahalik, N. P., Nambiar, A. N., 2010. "Trends in food packaging and manufacturing systems and technology". *Trends in Food Science & Technology*, **21**, pp. 117–128.
- [3] <http://www.wrapping-machines.co.uk>
- [4] <http://www.tosa.it>
- [5] <http://www.areapackaging.com>
- [6] Belluzzi, O., 1989. "Scienza delle costruzioni 1", Vol. 1 della Serie *Scienza delle Costruzioni*. Zanichelli, Bologna, p. 657.
- [7] Viola, E., 1993. "Esercitazioni di scienza delle costruzioni. 1". Vol. 1 della Serie *Esercitazioni di Scienza delle Costruzioni*. Pitagora, Bologna.
- [8] UNI EN 10025-2, 2005. "Prodotti laminati a caldo di acciai per impieghi strutturali: Parte 2: Condizioni: tecniche di fornitura di acciai non legati per impieghi strutturali".
- [9] <http://www.atlantastretch.com/ita/omega.php>

ON SHAFT-HUB HYBRID COUPLINGS

Dario Croccolo

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: dario.croccolo@unibo.it*

Massimiliano De Agostinis

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: m.deagostinis@unibo.it*

Nicolò Vincenzi

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: nicolo.vincenzi@unibo.it*

Abstract. *Press-fitted and adhesively bonded joints (Hybrid Joints) are increasingly used as an alternative way to traditional structural joining techniques. The main achievable benefits can be summarized in the possibility of maximizing the load transfer (torque or axial) and reducing both the weight and the stress field of the components, by taking advantage of the adhesive strength. Hybrid joints studies can be found in technical literature mainly on steel-steel components (Steel Hybrid Joints). The aim of this paper is to provide some relevant information on the static and fatigue strength properties in the case of steel-aluminium components (Mixed Hybrid Joints), from the experimental tests performed on a high strength, single-component adhesive, which cures anaerobically. The use of the adhesive increases the press-fitted joint performances, with respect to its release force: the adhesive static shear strength is about 9MPa, whereas the adhesive endurance limit is about 6MPa, in presence of a stress ratio $R=0.1$.*

Keywords: *anaerobic adhesive, aluminium, joint design, fatigue, bonding.*

List of symbols:

A	coupling surface [mm ²]
$D_{ext,H}$	external hub diameter [mm]
$D_{int,H}$	internal hub diameter [mm]
D_C	coupling diameter [mm]
$D_{ext,S}$	external shaft diameter [mm]
$D_{int,S}$	internal shaft diameter [mm]
E	Young's modulus (shaft E_S , hub E_H) [MPa]
$F_{int,ad}$	interference, adhesive contribution to the axial release force of the joint [N]
$F_{coupling}$	force to join shaft and hub in press-fit [N]
L_C	coupling length [mm]
p_C	coupling pressure [MPa]
$Q_{S,H}$	ratio between the internal and external shaft, hub diameters [mm]
$R_{z,a}$	surface roughness [μ m]
U, Z	nominal, actual interference [mm]

ε_θ	hoop strain
μ_A	axial static coefficient of friction
ν	Poisson's ratio (shaft ν_S , hub ν_H) [MPa]
τ_{ad}	adhesive static shear strength [MPa]

1. INTRODUCTION

Press-fitted and adhesively bonded joints (*Hybrid Joints - HJ*) are increasingly used as an alternative way to traditional structural joining techniques. The main benefits provided by the use of *HJ* can be summarized in the possibility of maximizing the load transfer (torque or axial) and reducing both the weight and the stress field in the components, by taking advantage of the adhesive strength: the adhesive can either replace traditional joining techniques (i.e. keys, pins, interference-fits, bolted joints...) or can actually improve the overall performances if used in combination with traditional techniques. As a matter of fact, dry press-fitted couplings achieve only 20-30% of nominal surface contact (metal to metal), whereas the adhesive presence is able to fill the voids (air), reaching 100% of contact area [1]. However, the strength of *HJ* must be evaluated accurately, since it is affected by various factors, such as the coupling pressure [2], the type and the way of assembling [3,4], the type of materials in contact [5,6], the surface roughness [7], the curing type and the curing methodology [8], the operating temperature [9] and the loading type [10]. Most of technical literature about *HJ* is mainly focused on steel-steel components (*Steel Hybrid Joints - SHJ*) [2-5,11,12] and on their verification technique (non-destructive testing) [13,14]. The use of different materials implies the diffusion of new joining technologies. For example, the automotive industry is strongly developing lightweight and energy efficient vehicles [15-17]: efforts are mainly directed towards the substitution of steel by aluminium or other lightweight materials. Therefore, this paper aims at providing some relevant information on static and fatigue strength properties in case of steel-aluminium components (*Mixed Hybrid Joints - MHJ*) which are press-fitted and bonded via a high strength, single-component adhesive that cures in absence of oxygen (Loctite 648 [18]). *MHJ* can be realized by an aluminium shaft coupled with a steel hub (aluminium-steel), or by a steel shaft coupled with an aluminium hub (steel-aluminium). The external forces (coupling, release or fatigue loads) are always applied to the shaft.

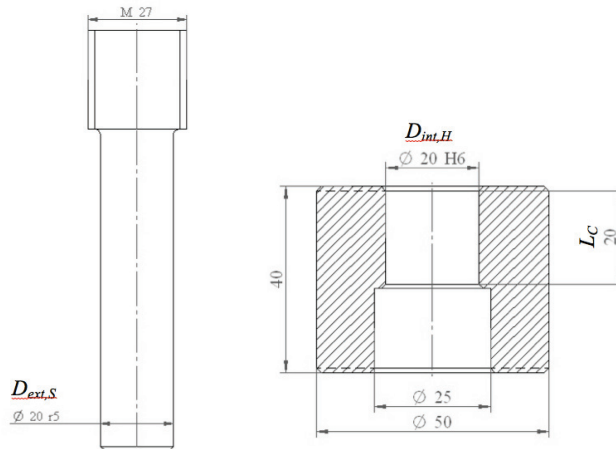


Figure 1. Geometrical dimensions of the specimens in millimeters: shaft made of steel (left) and hub made of aluminium (right).

2. MATERIALS AND METHOD

Experimental tests were performed on steel shafts made of 39NiCrMo3 [19] and aluminium hubs made of EN AW-6082T6 [20]. An example of the components and their geometrical dimensions is reported in Fig. 1, while their material properties, obtained by direct testing according to UNI-EN 10002, are reported in Tab. (1); specimens were obtained by means of a CNC lathe (T120 Meccanica Cortini, Forlì, ITALY).

The fundamental design parameter of press-fitted couplings is the axial release force F_{int} [N], reported in Eqn. (1), which depends on the axial static friction coefficient μ_A , coupling pressure p_C [MPa] and the contact surface A [mm²].

$$F_{int} = \mu_A \cdot p_C \cdot A \quad (1)$$

The axial static friction coefficient μ_A strictly depends on the coupled materials and the surface conditions of the components: μ_A is usually within the range 0.04-0.7 [21]. Thus, it must be accurately evaluated in order to reduce the variation in the F_{int} calculation. The coupling pressure p_C , which is generated by the actual interference Z [mm] between the shaft and the hub, is a function of the stiffness of the coupled components, as highlighted in Eqn. (2) [22,23].

$$p_C = \frac{Z/D_C}{\frac{1}{E_H} \cdot \left(\frac{1+Q_H^2}{1-Q_H^2} + \nu_H \right) + \frac{1}{E_S} \cdot \left(\frac{1+Q_S^2}{1-Q_S^2} - \nu_S \right)} \quad (2)$$

D_C [mm] is the coupling diameter, equal to the mean of the external shaft $D_{ext,S}$ [mm] and the internal hub $D_{int,H}$ [mm] diameters, E [MPa] and ν are the Young's modulus and the Poisson's ratio, respectively (hub: E_H , ν_H ; shaft: E_S , ν_S) and Q is the ratio between the internal and external diameters of the hub ($Q_H = D_{int,H}/D_{ext,H}$) and of the shaft ($Q_S = D_{int,S}/D_{ext,S}$). The actual interference Z [mm] is lower than the nominal one U [mm] when the assembly is realised by means of a hydraulic press with the shaft press-fitted into the hub and both set at the same temperature. If the stiffness of the parts has a similar value (for example, in case of the same material type), the difference between U and Z can be related to the surface roughness R_a or R_z [mm] of the components (hub: $R_{a-z,H}$; shaft: $R_{a-z,S}$), as proposed in [24,25] and reported in Eqn. (3).

$$Z = \begin{cases} (D_{ext,S} - D_{int,H}) - 0.8 \cdot (R_{z,H} + R_{z,S}) \\ (D_{ext,S} - D_{int,H}) - 3 \cdot (R_{a,H} + R_{a,S}) \end{cases} \quad (3)$$

The relationships reported in Eqn. (3) have been obtained empirically from several tests performed on steel-steel assemblies. In order to evaluate the actual coupling pressure accurately, especially in presence of different coupled materials, it is advisable to apply a strain gage on the external surface of the hub, which is able to provide the external hoop strain ε_θ , as reported in Fig. 2. In light of this, ε_θ can be related to p_C by applying the thick walled cylindrical theory reported in Eqn. (4), and well demonstrated by Croccolo et al. [26].

Table 1 Material properties of the specimens used as evaluated according to UNI EN 10002

MATERIAL	E [GPa] Young's modulus	ν Poisson's ratio	S_Y [MPa] Yield stress	S_U [MPa] Ultimate stress	A% Elongation
<i>Steel</i> 39NiCrMo3	207	0.29	900	1050	3
<i>Aluminium</i> 6082 T6	69	0.33	304	343	9

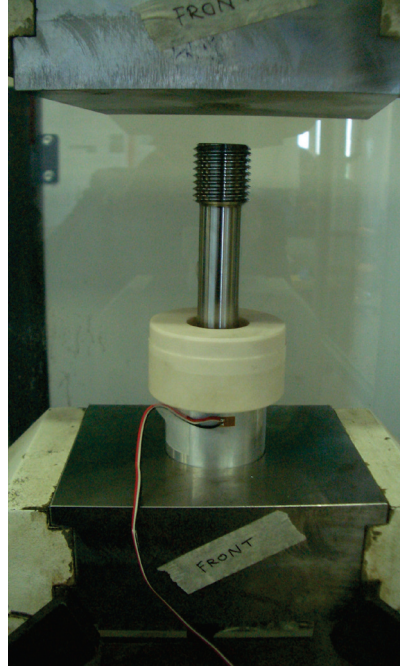


Figure 2. Example of the joint coupling phase with the strain gage applied on the external surface of the hub (quarter-bridge connection in Wheatstone circuit).

$$P_C = \frac{E_H \cdot \varepsilon_\theta \cdot (1 - Q_H^2)}{2 \cdot Q_H^2} \quad (4)$$

Finally, the contact surface A is a function of the coupling diameter D_C and of the coupling length L_C [mm]: $A = \pi \cdot D_C \cdot L_C$.

The production of HJ begins with the spreading of the adhesive on the female component (the hub), followed by the assembly operation performed with the same procedure as used for press-fitted couplings. The mechanical interference and the adhesive strength act as parallel connected springs [12] and, therefore, the axial release force of HJ can be evaluated by applying the superposition of these effects, as shown in Eqn. (5): the two contributions (interference and adhesive) can be calculated separately as described in the next sections.

Table 2 Summary of results for steel-aluminium joints purely press-fitted

<i>Purely press-fitted joints</i>	<i>U</i> [mm]	<i>p_C</i> [MPa]	<i>R_{a,H} before</i> [μm]	<i>R_{a,S} before</i> [μm]	<i>R_{a,H} after</i> [μm]	<i>R_{a,S} after</i> [μm]	<i>F_{coupling}</i> [kN]	<i>F_{int}</i> [kN]	<i>μ_A</i>
1	0.026	40.0	2.44	0.99	0.30	1.03	26.50	26.00	0.52
2	0.016	26.1	2.44	0.92	0.34	1.05	16.50	14.00	0.43
3	0.016	27.2	2.44	0.89	0.69	0.90	17.90	15.00	0.44
4	0.019	35.5	2.44	1.00	0.51	1.15	23.00	21.00	0.47
5	0.016	25.3	2.45	0.97	0.40	1.10	15.50	13.00	0.41
6	0.018	32.1	2.44	0.91	0.37	1.01	22.00	19.50	0.48
7	0.021	33.9	2.44	0.95	0.32	0.97	22.60	20.00	0.47
8	0.024	38.1	2.45	0.94	0.34	1.02	26.60	23.50	0.49
Mean value	0.020	32.3	2.44	0.95	0.41	1.03	21.30	19.00	0.46

$$F_{tot} = F_{int} + F_{ad} = (\mu_A \cdot p_C + \tau_{ad}) \cdot A \quad (5)$$

The friction coefficient μ_A is considered constant, both in presence and in absence of adhesive, because metals remain in contact (without any adhesive interposition) in the peaks of the roughness whereas the liquid adhesive fills the voids of all depressions [3]. The adhesive contribution F_{ad} [N] depends on the coupling surface A and the mean adhesive static strength τ_{ad} [MPa], which must be accurately calculated.

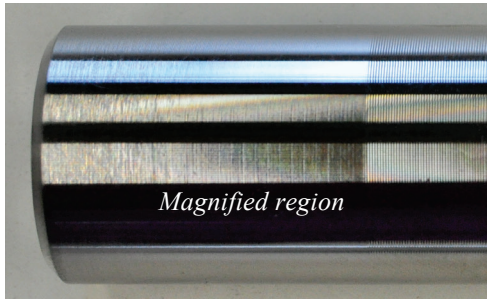
3. RESULTS

Mixed hybrid joints (MHJ): static strength

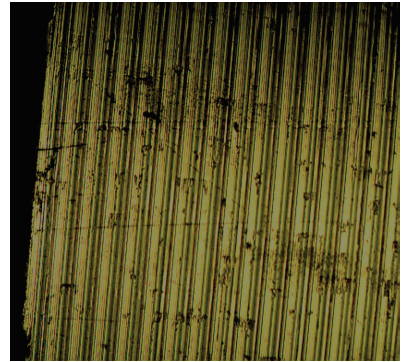
Three sets of tests were performed statically in order to evaluate *MHJ* performances in comparison to purely press-fitted joints: (i) dry press-fitted, (ii) adhesively bonded slip-fitted and (iii) adhesively bonded press-fitted.

Dry press-fitted joints. First, 8 different coupling and release tests were performed on a 100 kN hydraulic press with the purpose of evaluating the static strength (axial release force, F_{int}) of the pure press-fitted steel-aluminium joints (Fig. 1). The amount of nominal interference U is governed by the ISO system of limits and fits [27]: the tolerance amount $\phi 20H6/r5$ (Fig. 1) is able to provide the range $U_{min-max}=0.013\text{--}0.037$ mm. The actual coupling pressure p_C was evaluated according to Eqn. (4). Finally, the axial static friction coefficient μ_A was calculated by applying Eqn. (6), in which F_{int} is given by the load cell of the hydraulic press at the initial release (movement) of the specimens, and the coupling surface A is equal to 1.257 mm^2 . Results are reported in Tab. (2).

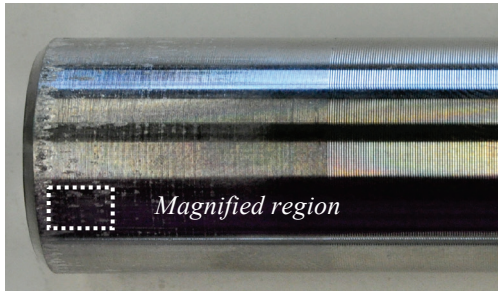
$$\mu_A = \frac{F_{int}}{p_C \cdot A} \quad (6)$$



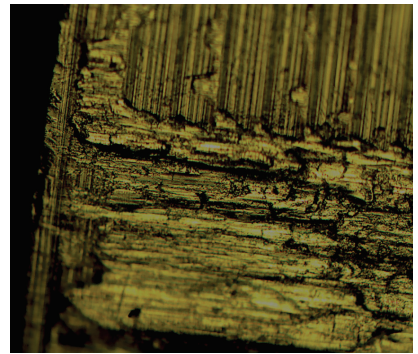
(a,1)



(a,2)



(b,1)



(b,2)

Figure 3. Pictures of shaft surfaces before coupling (a) and after release (b) operations: (1) 1x magnification, (2) 25x magnification.

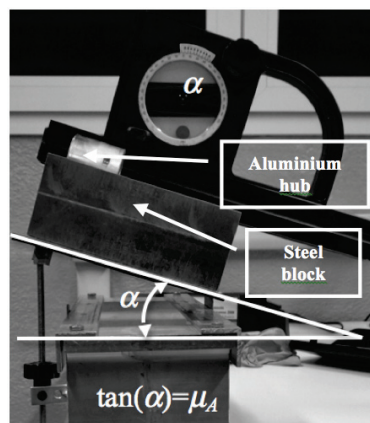


Figure 4. Example of sliding test: experimental evaluation of friction angle α between the shaft and hub materials.

It is possible to highlight that the actual interference Z depends only on the surface roughness variation of the aluminium hub, before ($R_{a,H \text{ before}}$) and after ($R_{a,H \text{ after}}$) the coupling. Examples of shaft surfaces, before and after the coupling, are shown in Fig. 3. It is possible to identify some regions where aluminium dust has been transferred from the hub to the shaft. This occurrence is magnified by the different stiffness of coupled materials (aluminium and steel), which produces a high wear of aluminium during the press-fit operation [28]. As a matter of fact, the axial force needed to join the parts ($F_{coupling}$) is always higher than the release one (F_{int}). The mean value of the axial static friction coefficient μ_A is equal to 0.46, with a standard deviation equal to 0.03. Some further sliding tests, on the same studied materials, performed on a skew plane have confirmed such value (Fig. 4).

Adhesively bonded slip-fitted joints. A second set of 8 tests was performed on slip-fitted couplings, which had a clearance between the steel shaft and the aluminium hub. These tests allow evaluating the mean strength of the purely anaerobic adhesive (τ_{ad}), without the interference contribution. The clearance value (difference between diameters) was realised in the range 0.02-0.05 mm since this value is associated with the maximum strength for *SHJ*, as showed by Sekercioglu et al. [4]. Specimens were cleaned with Loctite 7061, which is a general purpose cleaner for preparing surfaces to be bonded with adhesive, then, specimens were cured for 24 h at 40°C. All tested specimens showed an adhesion failure, like that reported in Fig. 5. The adhesive always remained on the steel adherend. This occurrence is mainly due to the lower value of surface energy of aluminium (wettability) with respect to steel, as well demonstrated by Borsellino et al. [29] and Comyn [30] by the determination of the contact angles at solid/liquid interfaces. The mean shear strength τ_{ad} of the adhesive for slip-fitted steel-aluminium joints was 9.6 MPa with a standard deviation equal to 0.2 MPa. Results are reported in Tab. (3). A comparison between slip-fitted steel-steel joints [4] and slip-fitted aluminium-steel joints is reported in Fig. 6.

Adhesively bonded press-fitted joints. Finally, a third set of 20 different tests was performed on press-fitted and adhesively bonded steel-aluminium joints, with the purpose of evaluating the total strength of *MHJ*, compared to the previous ones. The results plotted in Fig. 7 highlight the increase in scatter for *MHJ* with respect to the two previous sets of tests. The static strength of the adhesive F_{ad} and the adhesive mean shear strength τ_{ad} were derived from Eqn. (7), in which the interference contribution (F_{int}) is given by applying Eqn. (1), and the total axial decoupling force (F_{tot}) is given by the force transducer (load cell 661 MTS, Eden Prairie MN, USA) of the hydraulic press.

$$F_{ad} = F_{tot} - F_{int} \Rightarrow \tau_{ad} = \frac{F_{tot} - F_{int}}{A} = \frac{F_{tot}}{A} - \mu_A \cdot p_C \quad (7)$$

Since the coupling is realized by a press-fit, the variation of release force (Fig. 7) can be attributed to the amount of adhesive that actually remains at the interface of coupled surfaces. Accordingly, the frictional contribution was considered constant.

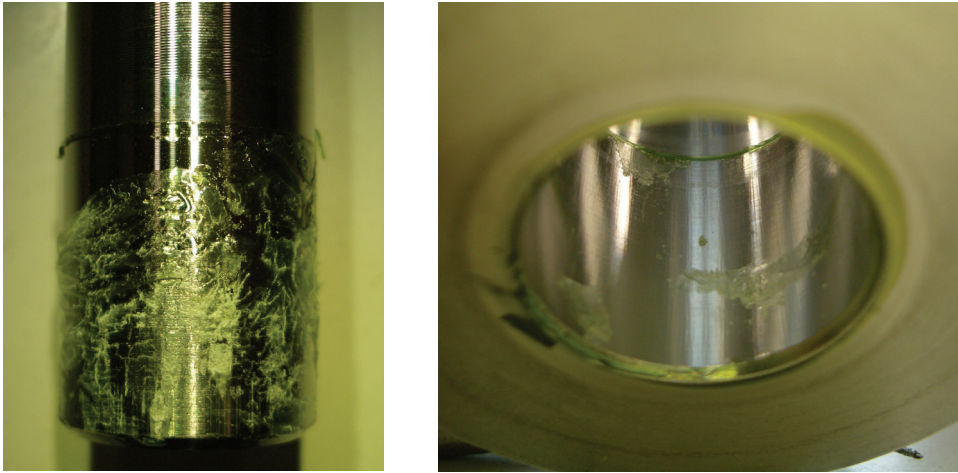


Figure 5. Example of adhesive failure in slip-fitted aluminium-steel joint: the adhesive is nearly absent on the aluminium surface.



Figure 6. Mean adhesive shear strength in case of slip-fitted joints (steel-steel data from [4], steel-aluminium data from the present work).

Table 3 Summary of results for steel-aluminium adhesively bonded slip-fitted joints

<i>Adheisvely bonded slip-fitted joints</i>	$D_{ext,S}$ [mm]	$D_{int,H}$ [MPa]	<i>Gap</i> [mm]	L_c [mm]	F_{ad} [kN]	τ_{ad} [MPa]
1	20.02	19.97	0.05	20	12.20	9.7
2	20.01	19.96	0.05	20	11.80	9.4
3	20.02	19.97	0.05	20	12.20	9.7
4	20.00	19.95	0.05	20	11.70	9.3
5	20.01	19.96	0.05	20	12.10	9.6
6	20.01	19.96	0.05	20	12.30	9.8
7	20.02	19.97	0.05	20	11.90	9.5
8	20.00	19.95	0.05	20	12.30	9.8
Mean value	20.01	19.96	0.05	20.00	12.06	9.6

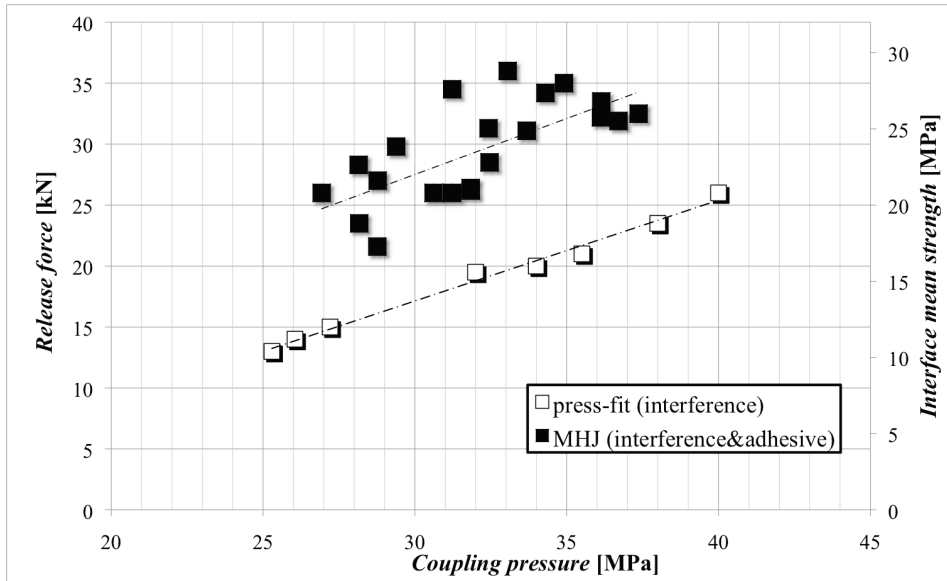
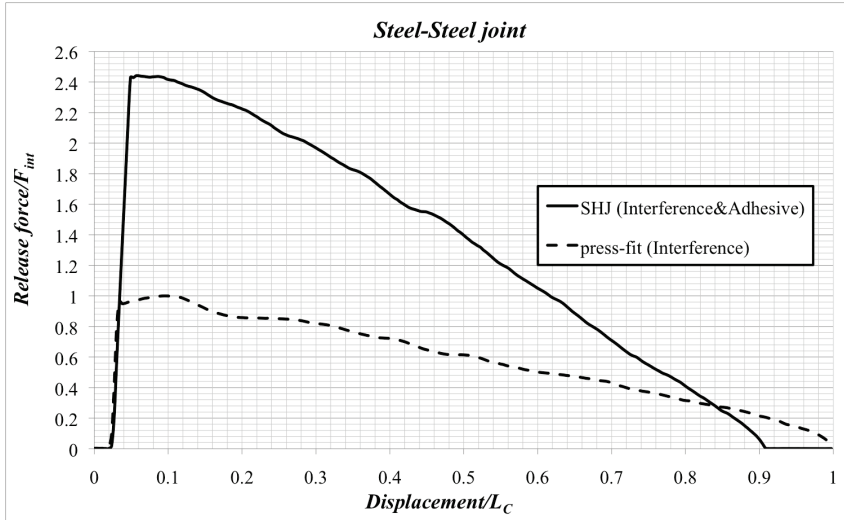


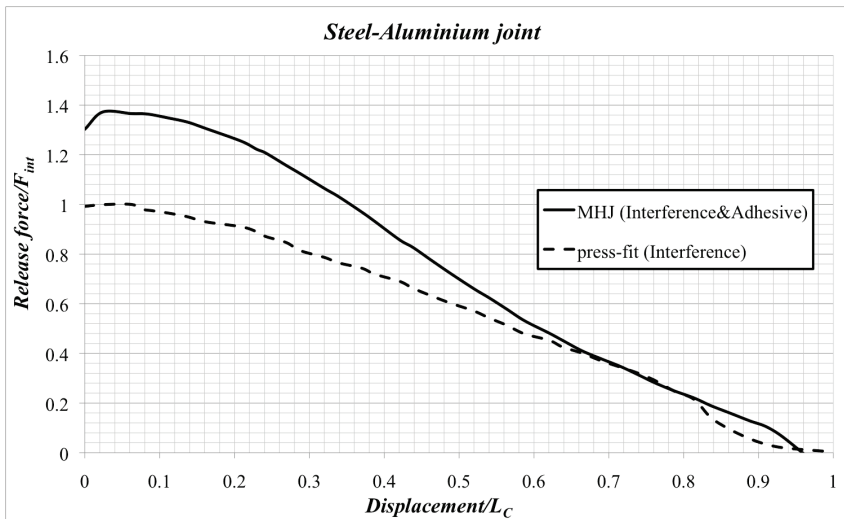
Figure 7. Comparison between purely press-fitted and *MHJ*: axial release force and interface mean strength as a function of the coupling pressure.

The overall performance of the joint increases both in case of *SHJ*, as demonstrated in [1-8,12-14,26], and in case of *MHJ* as highlighted in this work. The axial release force and the interface mean strength are plotted in Fig. 7 as a function of the actual coupling pressure. In the case of purely press-fitted joints, the interface mean strength is equal to $\mu_A p_C$ whereas in case of *HJ* it is equal to $\mu_A p_C + \tau_{ad}$. The adhesive mean strength τ_{ad} is within the range of 6-12MPa, while its average value is equal to 9.2MPa with a standard deviation of 1.9MPa. Results are reported in Tab. (4). The average value of τ_{ad} in press-fitted and adhesively bonded joints is very close to that found for slip-fitted joints (9.6 MPa), even if the standard deviation increases (1.9 MPa vs 0.2 MPa). An example of the force trend plotted as a function of the displacement during the release phase is reported in Fig. 8. Values, for both *SHJ* [26] (Fig. 8a) and for *MHJ* (Fig. 8b), have been conveniently

made dimensionless by normalizing them to the interference release force (F_{int}) and to the coupling length (L_c). In case of *MHJ*, the total and the adhesive strengths have been noticed to be strictly related to the maximum assembly force (coupling force) as reported in Fig. 9 and in Fig. 10, respectively. In the light of this, it is possible to define two constants, K_1 and K_2 , calculated by the ratio between the total strength and the coupling load, and by the ratio between the adhesive strength and the coupling load, respectively (Eqn. 8). Their trends are reported in Fig. 11, as a function of the coupling pressure.



(a)



(b)

Figure 8. Release forces for *HJ* compared to the purely press-fitted ones: (a) *SHJ* [26] and (b) *MHJ*.

Table 4 Summary of results for steel-aluminium press-fitted adhesively bonded joints

<i>Adheisvely bonded press- fitted joints</i>	<i>U</i> [mm]	<i>F_{coupling}</i> [kN]	<i>F_{tot}</i> [kN]	<i>F_{ad}</i> [kN]	<i>τ_{ad}</i> [MPa]
1	0.018	25.00	31.30	12.87	10.2
2	0.015	22.30	28.30	12.30	9.8
3	0.015	19.00	26.00	10.69	8.5
4	0.016	26.90	29.80	13.10	10.4
5	0.018	33.00	36.00	15.01	11.9
6	0.019	28.00	35.00	14.96	11.9
7	0.016	15.60	21.60	8.55	6.8
8	0.020	29.10	31.90	11.04	8.8
9	0.016	22.50	27.00	10.65	8.5
10	0.017	23.00	26.40	8.61	6.8
11	0.020	31.00	33.50	12.96	10.3
12	0.019	30.60	34.20	14.71	11.7
13	0.017	17.50	26.00	8.60	6.8
14	0.018	25.00	31.10	11.96	9.5
15	0.018	26.00	28.50	10.05	8.0
16	0.017	21.00	26.00	8.75	7.0
17	0.015	18.00	23.50	8.02	6.4
18	0.020	25.80	32.20	11.67	9.3
19	0.017	30.20	34.50	14.30	11.4
20	0.020	25.30	32.50	11.26	9.0
Mean value	0.017	24.66	29.77	11.50	9.2

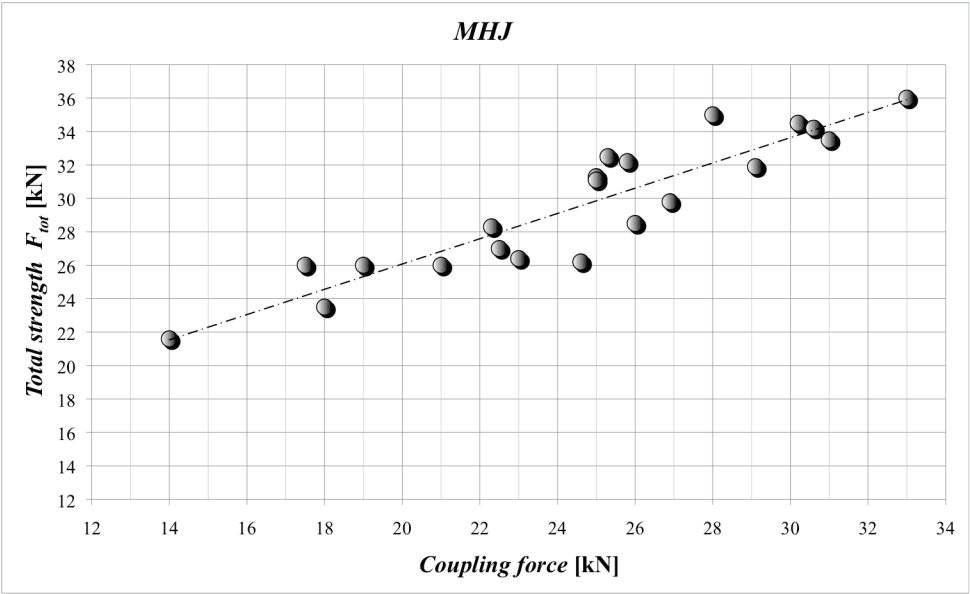


Figure 9. Total strength $F_{tot} = F_{int} + F_{ad}$ plotted as a function of the coupling force.

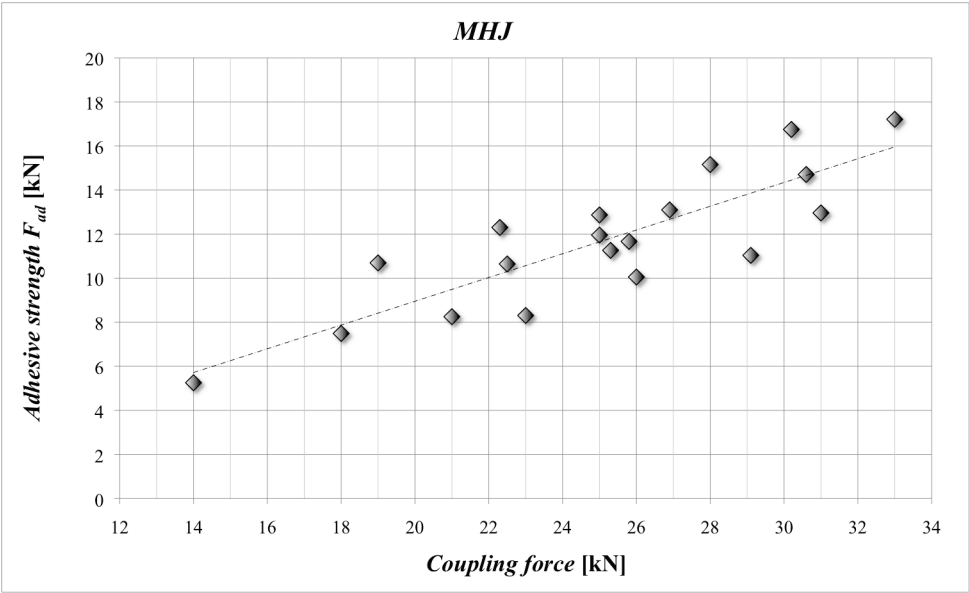


Figure 10. Adhesive strength contribution F_{ad} plotted as a function of the coupling force.

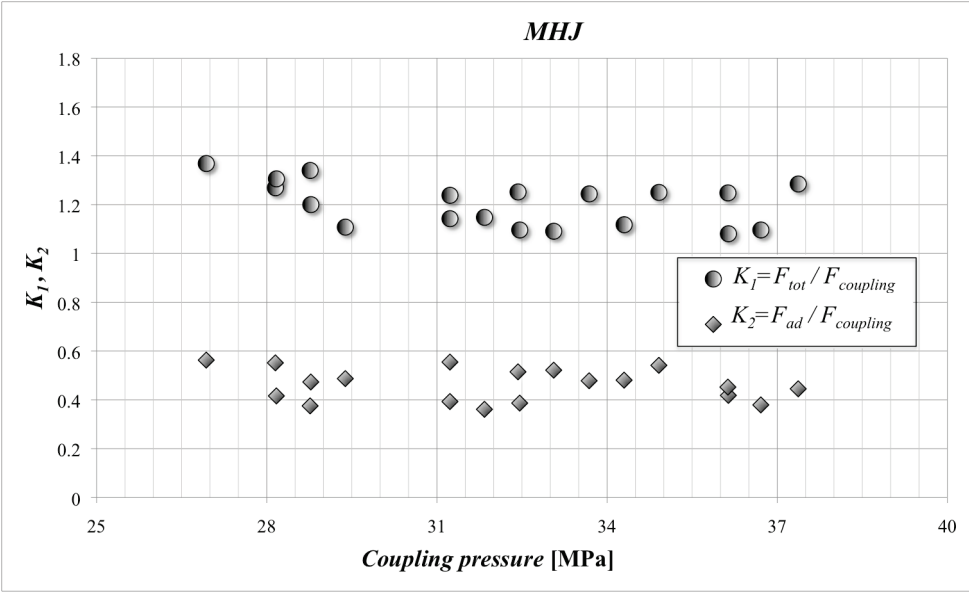


Figure 11. K_1 and K_2 trends as a function of the coupling pressure.

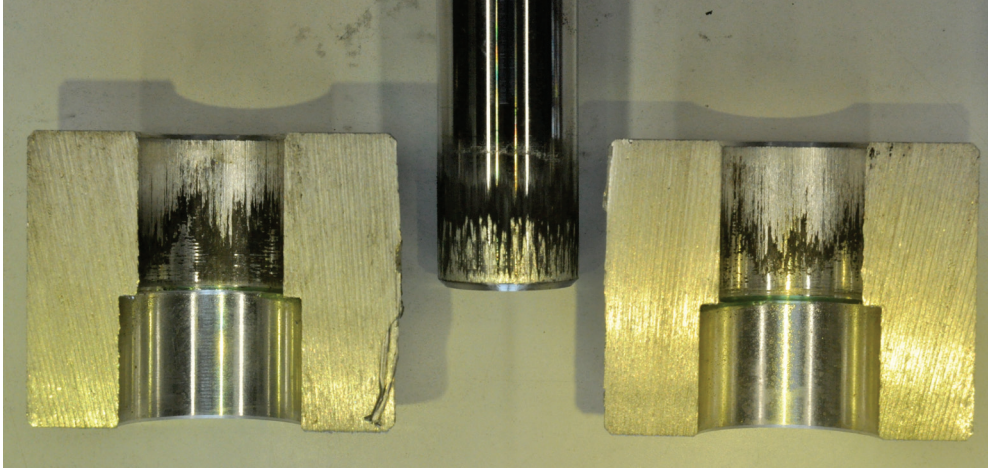


Figure 12. Example of fracture surfaces in *MHJ* (press-fitted and adhesively bonded): aluminium hub cut in halves and steel shaft.

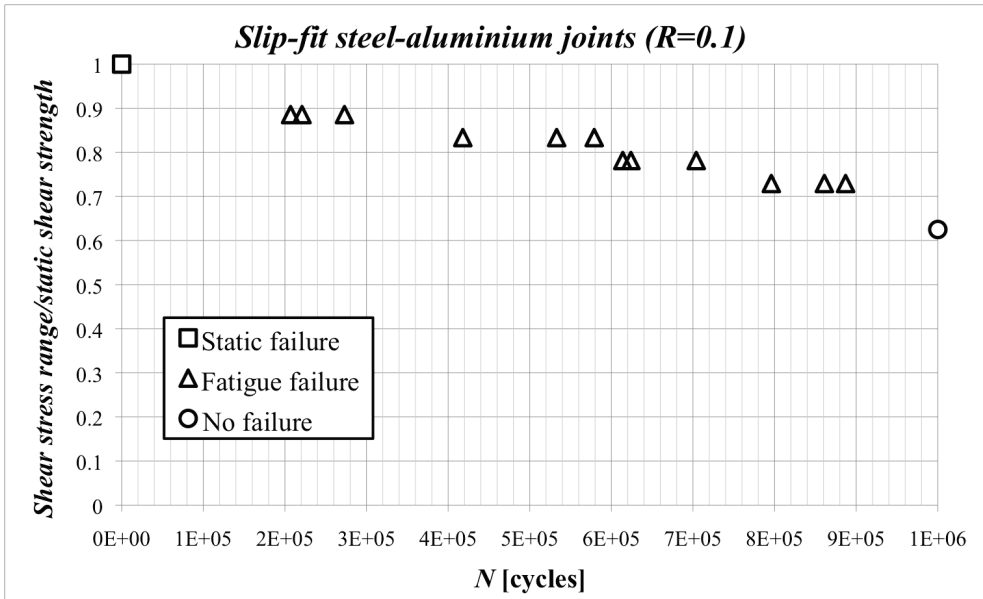


Figure 13. *S-N* diagram for slip-fitted steel-aluminium joints (stress ratio $R=0.1$).

$$\begin{cases} K_1 = \frac{F_{tot}}{F_{coupling}} = \frac{F_{int} + F_{ad}}{F_{coupling}} \approx 1.23 \\ K_2 = \frac{F_{ad}}{F_{coupling}} \approx 0.47 \end{cases} \quad (8)$$

Finally, an example of fracture surfaces for press-fitted and adhesively bonded steel-aluminium joints is reported in Fig. 12: the different aspect of the adhesive here, with respect to Fig. 5, is due to the interference (coupling pressure) between the components, in place of clearance (gap).

Mixed hybrid joints (MHJ): fatigue strength

A total of 15 fatigue tests based on stress-life ($S-N$) approach were performed on slip-fitted joints, with the purpose of evaluating the purely anaerobic adhesive fatigue strength [31]. The stress ratio R was set equal to 0.1 (nearly “pulsating tension” condition) while the test frequency f was set equal to 20 Hz. The shear stress range $\Delta\tau$ [MPa], dimensionless with respect to the static shear stress (9.6 MPa for slip-fitted steel-aluminium joints) is reported in Fig. 13, as a function of the number of cycles N . The ratio between the shear stress range and the static shear strength at 10^6 cycles results equal to 0.62 ($\Delta\tau_\infty \approx 0.62 \cdot \tau_{ad}$).

Then, fatigue results determined for slip-fitted steel-aluminium joints were applied to 6 different press-fitted steel-aluminium ones (with coupling pressures in the same range as of the ones tested statically), in order to verify the $S-N$ curve applicability (Fig. 13) to press-fitted joints. Remembering that the interference contribution ($F_{int} = \mu_A p_C A$) is constant with respect to fatigue, 3 joints were loaded as

$$\begin{cases} F_{\min} = 0.1 \cdot F_{\max} \\ F_{\max} = (\mu_A \cdot p_C + \Delta\tau_\infty) \cdot A = (\mu_A \cdot p_C + 0.62 \cdot \tau_{ad}) \cdot A \end{cases} \quad (9)$$

While other 3 joints as

$$\begin{cases} F_{\min} = 0.1 \cdot F_{\max} \\ F_{\max} = (\mu_A \cdot p_C + 1.3 \cdot \tau_\infty) \cdot A \end{cases} \quad (10)$$

This test method was used because the coupling pressure cannot be exactly constant, due to the tolerance differences in the interference values.

Specimens loaded by the fatigue cycle reported in Eqn. (9) reached 10^6 cycles, without disengaging, whereas specimens stressed by the fatigue cycle reported in Eqn. (10), beyond the expected fatigue limit, lasted between 350,000 and 550,000 cycles. The $S-N$ curve reported in Fig. 13 is, therefore, suitable also for the fatigue strength prediction of the adhesive contribution in press-fitted and adhesively bonded steel-aluminium joints.

4. DISCUSSION

MHJ with aluminium shaft and steel hub (aluminium-steel) had several problems due to the high elastic deformation of aluminium that modifies the radial interference when loaded both in tension and in compression [26]. In tension, the actual interference decreases and a peel stress occurs on the adhesive, so that the joint performances decrease dramatically; whereas in compression the actual interference increases, resulting in an overestimation of the joint strength. *MHJ* with steel shaft and aluminium hub (steel-aluminium) are, on the contrary, able to overcome the aforementioned issues and, therefore, they are the most suitable to be used.

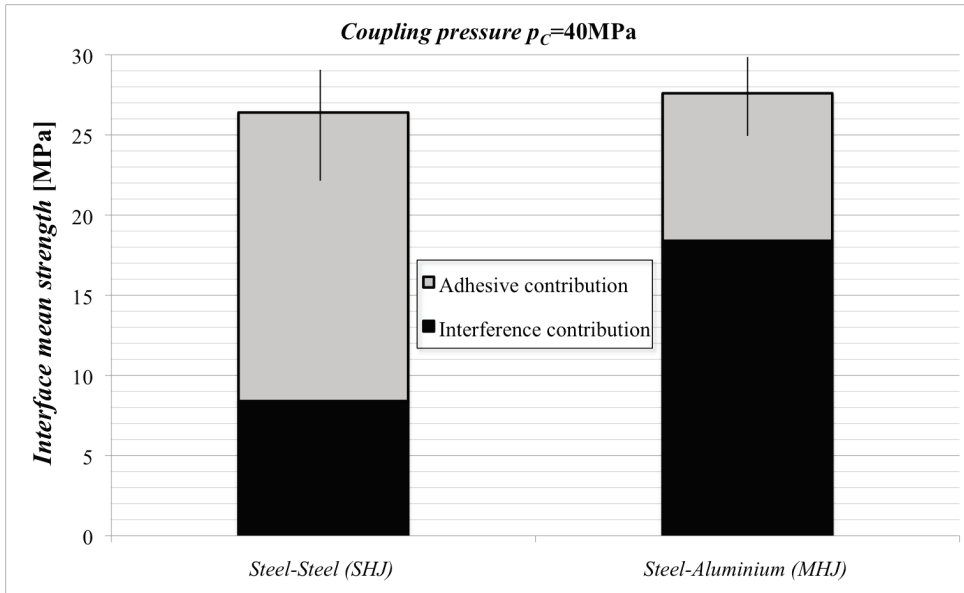


Figure 14. Comparison of total interface mean strength between *SHJ* [26] and *MHJ*: interference contribution ($\mu_A p_C$) vs adhesive contribution (τ_{ad}).

The anaerobic adhesive increases the joint performances both in *SHJ* and in *MHJ* with steel shaft and aluminium hub. In press-fitted joints, by considering p_C within the 20-40 MPa range (standard coupling pressure values), the adhesive contribution in *SHJ* is higher than the interference contribution (up to 250%) [26], whereas the adhesive contribution in *MHJ* is lower than the interference contribution (about 40%); therefore, the effectiveness of anaerobic adhesive decreases in presence of the aluminium adherend. However, the same coupling pressure produces a similar total strength of the joint, since the friction coefficient in *MHJ* ($\mu_A = 0.46$) is nearly 220% the *SHJ* one ($\mu_A = 0.21$ [26]). This occurrence is shown in Fig. 14: in *SHJ* the adhesive is responsible for about 65% of the total strength (interference produces only 35%), while in the *MHJ* the percentages are swapped (interference 64%, adhesive 36%).

K_1 and K_2 are very useful parameters to predict the *MHJ* static strength and its endurance limit, once the maximum coupling load is known (for instance derived by the load cell of the hydraulic press during the assembly phase).

Finally, the fatigue tests performed on *MHJ* (nearly pulsating tension, $R = 0.1$) highlighted that in presence of a shear stress range equal to 62% of the static shear strength, the number of cycles equal to 10^6 can be reached, for both slip-fitted and press-fitted hybrid joints. In short, in order to reach 10^6 cycles, the adhesive can be stressed up to 6 MPa, with a pulsating tension. In steel-steel press-fitted and adhesively bonded joints, Croccolo et al. [26] found that 10^6 cycles correspond to a shear stress range equal to 35% of the static shear strength ($\tau_{ad} \approx 18$ MPa). Therefore, the value of 6 MPa, regardless of the adherend involved, seems to be the endurance limit for the anaerobic adhesive studied (10^6 cycles, $R = 0.1$, according to [31]). In brief, the adhesive static shear strength changes with the coupled materials (*SHJ* and *MHJ*) because of the change in the type of failure, whereas the adhesive endurance limit remains the same.

5. CONCLUSIONS

Static and fatigue strength properties of mixed hybrid joints (steel-aluminium) were evaluated by performing several experimental tests on a high strength, single-component adhesive, which cures anaerobically. The adhesive contribution with respect to the total strength is restricted by the weak bonding forces between aluminium and adhesive, leading to adhesion failures. However, if coupling pressures of about 40 MPa (a standard value for interference-fit joints) are realized, steel-aluminium and steel-steel joints produce a similar total strength. In fact, the presence of higher friction coefficients is balanced by lower static adhesive strengths in steel-aluminium joints with respect to steel-steel ones. Furthermore, the endurance limit of mixed hybrid joints is also very close to that calculated for steel hybrid ones, at least within the analyzed ranges of pressure and in the presence of a high strength anaerobic adhesive. The good relationship between the maximum coupling force and the static adhesive strength has to be investigated in more detail in the future and, we believe, will be confirmed for different interference amounts and adhesive types.

Acknowledgements

The technical assistance of Eng. Paolo Proli (Director of Mechanical Engineering Laboratory) is gratefully acknowledged. A special acknowledgement is due to Dr. Piero Mauri, Henkel Loctite Italy, for his kind cooperation.

REFERENCES

- [1] D. Croccolo, R. Cuppini and N. Vincenzi, 2008. *Proceedings of the 10th International Conference on the Science and Technology of Adhesion and Adhesives*, St. Catherine College, Oxford, UK, pp. 277-81.
- [2] G. Bartolozzi, D. Croccolo and M. Chiapparini, 1999. *Österr. Ing. Z.* 5, 198-201.
- [3] E. Dragoni and P. Mauri, 2000. *Intl J. Adhesion Adhesives* 20, 315-21.
- [4] T. Sekercioglu, A. Gulsoz and H. Rende, 2005. *Mater. Design* 26, 377-81.
- [5] T. Sekercioglu, 2005. *Intl J. Adhesion Adhesives* 25, 352-57.
- [6] T. Sekercioglu and C. Meran, 2004. *Mater. Design* 25, 171-75.
- [7] T. Sekercioglu, H. Rende, A. Gulsoz and C. Meran, 2003. *J. Mater. Proc. Technol.* 142, 82-86.
- [8] R. Mengel, J. Haberle and M. Schlimmer, 2007. *Intl J. Adhesion Adhesives* 27, 568-73.
- [9] L. F. M. da Silva and R. D. Adams, 2007. *Intl J. Adhesion Adhesives* 27, 362-79.
- [10] R. D. Adams, J. Comyn and W. C. Wake, 1997. *Structural Adhesive Joints in Engineering*. Chapman & Hall, London.
- [11] H. Kawamura, T. Sawa, M. Yoneno and T. Nakamura, 2003. *Intl J. Adhesion Adhesives* 23, 131-40.
- [12] G. Romanos, 1999. *Intl J. Mater. Prod. Technol.* 14, 430-43.
- [13] D. Croccolo and R. Cuppini, 2006. *Intl J. Adhesion Adhesives* 26, 490-97.
- [14] D. Croccolo and R. Cuppini, 2009.. *Intl J. Adhesion Adhesives* 29, 280-85.
- [15] S. Birch, 1994. *Automotive. Eng.* 120, 8-12.
- [16] JJ. Robinson, 2003. *J. Mater.* 55, 6-10.

- [17] D. Croccolo, M. De Agostinis and N. Vincenzi, 2010. *Eng. Failure. Anal.* 17, 1173-87.
- [18] LOCTITE 1995. *Worldwide Design Handbook*.
- [19] UNI EN10083, 2006. Steels for quenching and tempering - Part 3: Technical delivery conditions for alloy steels.
- [20] UNI EN 573, 2005. Aluminium and aluminium alloys - Chemical composition and form of wrought products - Part 1: Numerical designation system.
- [21] D. Croccolo, R. Cuppini and N. Vincenzi, 2008. *Strain* 44, 170-79.
- [22] G. Niemann, H. Winter and B. R. Hohn, 2005. *Maschinenelemente Bd I*, Springer-Verlag, Berlin.
- [23] D. Croccolo and N. Vincenzi, 2009. *Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci.* 223, 2231-39.
- [24] DIN 7190, 2001. Interference fit – calculation and design rules.
- [25] NF E22-620, 1984. Assemblage frettes sur portee cylindrique, AFNOR, Paris la Defense.
- [26] D. Croccolo, M. De Agostinis and N. Vincenzi, 2010. *Intl J. Adhesion Adhesives* 30, 359-66.
- [27] UNI EN 20286, ISO system of limits and fits, 1995. Bases of tolerances, deviations and fits.
- [28] F. G. Kollmann, *Welle-Nabe Verbindungen*, Springer-Verlag, Berlin (1984).
- [29] C. Borsellino, G. Di Bella and V. F. Ruisi, 2009. *Intl J. Adhesion Adhesives* 29, 36-44.
- [30] J. Comyn, 1992. *Intl J. Adhesion Adhesives* 12, 45-49.
- [31] ISO 9664, 1993. Adhesives – Test method for fatigue properties of structural adhesives in tensile shear.

VALUTAZIONE DELLA CEDEVOLEZZA SELETTIVA IN GIUNTI ELASTICI PER APPLICAZIONI ROBOTICHE

Giovanni Berselli

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia.

Università di Bologna.

E-mail: giovanni.berselli@unibo.it

Marco Piccinini

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia.

Università di Bologna.

E-mail: marco.piccinini@mail.ing.unibo.it

Gabriele Vassura

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia.

Università di Bologna.

E-mail: gabriele.vassura@unibo.it

Sommario. *Si presenta una metodologia per quantificare le prestazioni di giunti elastici per applicazioni robotiche. In particolare, si considerano le caratteristiche di “cedevolezza selettiva”, riguardanti la capacità del giunto di fornire bassi valori di rigidità lungo una direzione desiderata conservando elevata rigidità in ogni altra direzione. In quest’ambito, si propongono indici di qualità utili a stabilire quale soluzione progettuale soddisfi al meglio i requisiti di una data applicazione. L’approccio è validato comparando quattro morfologie di giunto adatte a grandi spostamenti.*

Parole chiave: *giunti elastici per grandi spostamenti, applicazioni robotiche.*

1. INTRODUZIONE

Si definisce Giunto Elastico (GE) un tratto flessibile che collega due parti rigide e, deformandosi, ne consente il moto relativo. In accordo con [1], i GE possono essere utilizzati per sostituire le coppie cinematiche tradizionali (formate, ad esempio, mediante interposizione di cuscinetti volventi) all’interno di meccanismi a membri rigidi. Le strutture articolate così ottenute vengono denominate Meccanismi *Compliant* (Cedevoli) a cedevolezza concentrata (MC) e sono caratterizzati, come sottolineato dalla terminologia, da parti altamente deformabili concentrate in piccole regioni di spazio (nodi). I benefici dei GE, se comparati con le suddette coppie tradizionali, includono l’assenza di usure, giochi e attriti, nonché possibili

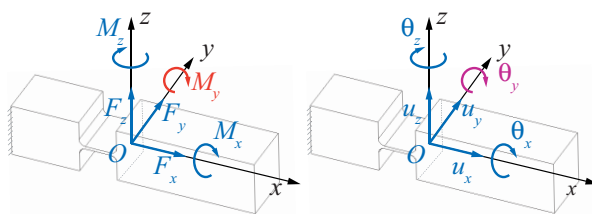


Figura 1. Carichi principali e secondari (sinistra) e spostamenti principali e secondari (destra).

riduzioni di pesi ed ingombri. Inoltre, l'avvento di nuovi materiali, nuove tecnologie (e.g. la prototipazione rapida) e nuovi campi di applicazione (e.g. i dispositivi MEMS), ha incentivato lo sviluppo di GE di classe superiore (e.g. [2, 3]).

In ambito robotico, l'introduzione di GE in sistemi articolati, come mani antropomorfe e protesi, può consentire la realizzazione di meccanismi snelli e leggeri che meglio si adattano all'ambizioso obiettivo di riprodurre le strutture biologiche di riferimento (e.g. [4, 5]). Tuttavia, questo tipo di applicazioni richiede l'utilizzo di GE compatibili con grandi spostamenti nonché elevate *cedevolezza selettive*, cioè bassa rigidezza lungo la direzione desiderata di movimento ed alte rigidezze lungo le altre direzioni [6, 7].

I vari tipi di GE proposti in letteratura sono classificabili in termini di entità di spostamento consentito ("piccoli spostamenti" [8] o "grandi spostamenti" [9]), numero e tipologia dei gradi di libertà (g.d.l.) (rotoidale o prismatico) [10], materiali utilizzati (multi- [4, 11, 12] o mono- materiale [13, 14]) e morfologia (e.g., GE di tipo "notch" o GE di tipo "leaf spring" [15]).

Con riferimento ai GE per grandi spostamenti, gli unici di reale interesse per le applicazioni di macro robotica, esistono diversi lavori recenti in cui si affronta il progetto e la modellazione di MC per mezzo di metodi agli elementi finiti (FEA) (e.g. [3]) o Modelli Pseudo-Rigidi (MPR) [9]. In quest'ultimo caso, il GE è modellato da una coppia rotoidale posta in parallelo ad una molla torsionale, in modo da semplificare la cinematica del giunto ad una rotazione nel piano.

Con riferimento ai GE di forma "semplice", come le travi snelle, sono note soluzioni analitiche in forma chiusa per spostamenti relativamente ampi [16–18] in grado di catturare gli effetti di svergolamento dovuti a carichi laterali. D'altro canto, il giunto ottimale per una determinata applicazione può essere caratterizzato da morfologie non banali, tali da non ammettere un trattamento mediante la teoria della trave.

In generale, la scelta di quale GE possa soddisfare al meglio la specifica applicazione dipende da numerosi fattori. Tuttavia, è possibile individuare alcune linee guida che possono indirizzare il processo di selezione di un generico GE, senza alcuna restrizione in termini di morfologia e tecniche di produzione del giunto stesso.

Giunti elastici per grandi spostamenti: concetti base

La maggior parte delle applicazioni richiede GE progettati per fornire un singolo g.d.l.. In questo caso, il GE è concepito per consentire uno *spostamento principale* lungo una direzione

di riferimento, qualora sottoposto ad un *carico principale* (coppia o forza), agente lungo la stessa direzione. Per esempio, la Fig. (1) mostra un *corner-filletted flexural hinge* (trave snella a sezione rettangolare con raccordi [8]), ovvero una struttura elastica che può agire come coppia rotoidale generando uno spostamento principale di rotazione θ_y sotto l'azione di un carico principale M_y . L'asse y è detto asse *compliant* [19, 20] ed il rapporto tra θ_y e M_y è chiamato *cedevolezza principale*.

Spostamenti secondari (anche chiamati effetti parassitici in [8]) lungo le altre direzioni di riferimento possono avere luogo nelle applicazioni reali per due ragioni: 1) presenza di *carichi secondari* agenti lungo le suddette direzioni; 2) presenza di spostamenti dell'asse *compliant*. In questo secondo caso, durante la deformazione del GE, l'asse compliant è soggetto a moto spaziale anche in assenza di un carico secondario. Il rapporto tra un qualsiasi spostamento secondario ed un qualsiasi carico viene definito *cedevolezza secondaria*.

E' bene notare che la presenza di carichi secondari è di solito difficilmente prevedibile poichè il MC potrebbe interagire con ambienti non strutturati. D'altro canto, le deformazioni dovute al carico principale (di solito fornito dal sistema di attuazione) possono essere valutate a priori tramite simulazione e considerate nel processo di design del dispositivo robotico. In ogni caso, trattandosi di moti indesiderati, gli spostamenti secondari devono essere evitati o, per quanto possibile, minimizzati.

Da questo punto di vista, per carichi dati, l'entità degli spostamenti secondari è strettamente dipendente dalla morfologia del GE: giunti che si comportano similmente, per quanto riguarda gli spostamenti principali, possono avere comportamenti piuttosto differenti in termini di spostamenti secondari. In particolare, come precedentemente riportato, la capacità di un GE di mantenere un'elevata cedevolezza lungo l'asse compliant ed un'elevata rigidità lungo ogni altra direzione è definita *cedevolezza selettiva* [8]. Per esempio, si considerino i GE per grandi spostamenti mostrati in Fig. (2). Tali GE sono caratterizzati da:

- Stesso intervallo di spostamenti (cioè una rotazione di $\pm 45^\circ$) al limite di snervamento;
- Stesso spostamento principale sotto l'applicazione dello stesso carico principale (cioè stessa cedevolezza principale, $C_{\theta_y M_y}$);
- Stesso materiale (i.e., stesso rapporto E/Y^1) e stessa tecnologia produttiva.

Un confronto critico di questi quattro GE consente di affermare che il giunto J_R (trave a sezione rettangolare, Fig. (2b)) è preferibile in applicazioni ad un singolo g.d.l. se confrontato con il giunto J_C , (trave a sezione circolare, Fig. (2a)), che è usualmente impiegato in MC spaziali [21]. Allo stesso modo, si può facilmente affermare che i giunti J_C e J_R hanno morfologie più semplici e dimensioni complessive maggiori se confrontati col giunto a spirale, *SPIR* (Fig. (2c)), e con il giunto elicoidale, *HEL* (Fig. (2d)), entrambe preferibili in termini di dimensioni qualora la complessità morfologica non sia un problema (per esempio, se il GE viene prodotto per mezzo di stereolitografia). Allo stesso modo, può essere difficoltoso scegliere tra *SPIR* e *HEL* senza uno studio approfondito del loro comportamento rispetto agli spostamenti secondari.

In risposta a tali quesiti, questa memoria propone un metodo di confronto per valutare la cedevolezza selettiva del GE sulla base di indici locali e globali. Il metodo è validato comparando i due GE la cui soluzione analitica è nota da letteratura, e cioè il giunto J_C ed il

¹ E indica il modulo di Young del materiale, Y indica la tensione di snervamento.

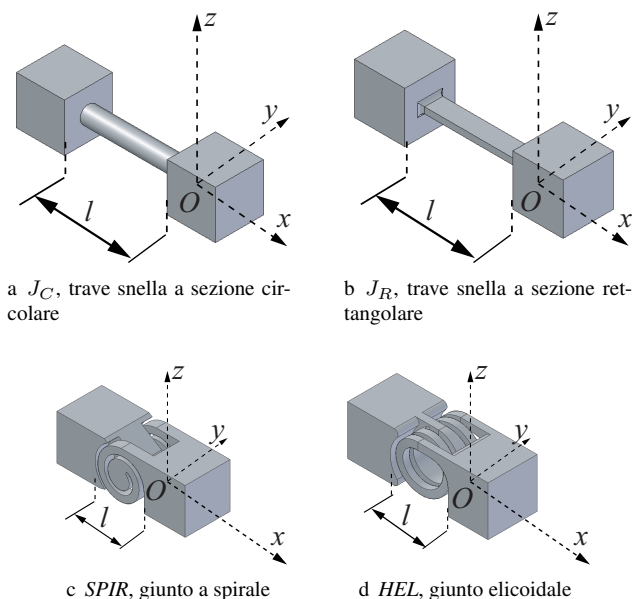


Figura 2. Giunti elastici di rotazione adatti a grandi spostamenti.

giunto J_R . Infine, il metodo è testato confrontando i giunti *SPIR* e *HEL*. In definitiva, si può affermare che il giunto *SPIR* è preferibile.

2. DESCRIZIONE DEL METODO

Allo scopo di confrontare differenti GE è necessario comprendere come i GE stessi reagiscano a carichi esterni. Nell'ambito di validità del principio di sovrapposizione degli effetti (materiali elastici lineari e piccole deformazioni), il comportamento cinematico di un GE nello spazio 3D può essere dedotto attraverso l'analisi della sua matrice di cedevolezza \mathbf{C} : dato il vettore dei carichi agenti in un punto (per esempio, il punto O in Fig. (1)), il vettore degli spostamenti lungo le direzioni di riferimento x, y, z è espresso da Eq. 1.

$$\Delta \mathbf{u} = \mathbf{C} \cdot \Delta \mathbf{f} \quad (1)$$

$$\Delta \mathbf{u} = \begin{bmatrix} \Delta u_x & \Delta u_y & \Delta u_z & \Delta \theta_x & \Delta \theta_y & \Delta \theta_z \end{bmatrix}^t$$

$$\Delta \mathbf{f} = \begin{bmatrix} \Delta f_x & \Delta f_y & \Delta f_z & \Delta m_x & \Delta m_y & \Delta m_z \end{bmatrix}^t$$

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix}$$

ove $\Delta \mathbf{u}$ è un vettore di spostamenti incrementali composti da 3 traslazioni incrementali (Δu_x , Δu_y , Δu_z) e altrettante rotazioni incrementali ($\Delta \theta_x$, $\Delta \theta_y$, $\Delta \theta_z$), $\Delta \mathbf{f}$ è una forza generalizzata “di disturbo” composta da tre forze incrementali (Δf_x , Δf_y , Δf_z) e altrettante coppie incrementali (Δm_x , Δm_y , Δm_z). Di conseguenza, \mathbf{C} C_{ij} è una matrice 6×6 con elementi di dimensioni fisiche non uniformi.

Una coppia rotoidale ideale, in grado di generare una rotazione pura anche in presenza di carichi secondari, presenterà una matrice \mathbf{C} in cui il solo termine $C_{55} = C_{\theta_y M_y}$, corrispondente alla rotazione lungo l’asse principale y sotto l’azione del carico principale M_y , è finito, essendo nulli gli altri termini. Al contrario un GE reale sarà caratterizzato da valori finiti dei coefficienti anche lungo le altre direzioni. In Fig. (3), viene mostrata una rappresentazione grafica qualitativa di matrici di cedevolezza \mathbf{C} per coppie rotoidali ideali e reali.

Tuttavia, essendo un operatore differenziale [22], la matrice di cedevolezza misura una proprietà locale che dipende dalla configurazione del GE. Pertanto, è chiaro che \mathbf{C} non può essere direttamente utilizzata come elemento di comparazione; in primis, poichè un’unica matrice non è sufficiente per descrivere globalmente il comportamento di un GE per grandi spostamenti. Secondariamente, poichè l’elevato numero di elementi che compongono \mathbf{C} non permette un confronto semplice ed intuitivo di differenti GE.

Pertanto, al fine di definire una metodologia di confronto, vengono proposti i seguenti passaggi concettuali:

1. Definizione e discretizzazione dello spazio di lavoro del GE;
2. Valutazione e normalizzazione della matrice di cedevolezza in ogni configurazione del giunto;

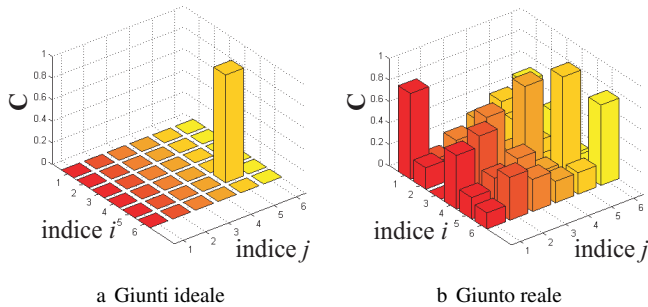


Figura 3. Rappresentazione grafica di matrici di cedevolezza per GE ideale (a) e reale (b). Indice i per le righe e indice j per le colonne.

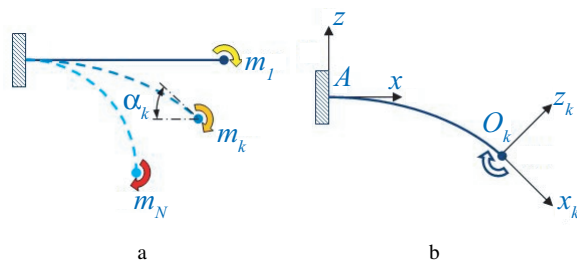


Figura 4. Discretizzazione dello spazio di lavoro (a) e sistemi di riferimento (b).

3. Definizione e valutazione di Indici di Qualità Locale (IQL), che caratterizzano la qualità del GE in termini di cedevolezza selettiva;
4. Definizione e valutazione di Indici di Qualità Globale (IQG) che riassumono la qualità complessiva del GE.

Definizione dello spazio di lavoro del giunto

Come detto, solo i GE caratterizzati da geometrie “semplici”, come le travi snelle sotto particolari condizioni di carico, possono essere studiati analiticamente nell’ipotesi di grandi spostamenti [16–18]. Pertanto, i GE caratterizzati da morfologie complesse (come *SPIR* o *HEL*) vengono analizzati per mezzo dei metodi FEA.

Si supponga che l’analisi di un numero finito di configurazioni del giunto, all’interno del suo stesso spazio di lavoro, sia sufficiente per ottenere una stima del comportamento del GE stesso rispetto alle proprietà di cedevolezza. Lo spazio di lavoro del GE, W_s , è dunque definito come *l’intervallo di spostamenti principali che può essere conseguito dal giunto prima di giungere a snervamento, qualora il GE sia soggetto unicamente ad un carico principale di intensità adeguata*. Lo spazio di lavoro è poi suddiviso in un insieme finito, N , di configurazioni del GE.

Per comprendere meglio questo approccio si può considerare il seguente esempio. In Fig. (4a) è rappresentato un GE a trave snella. Si definisca un sistema di riferimento inerziale con origine nel punto A ed un sistema di riferimento mobile con origine nel punto O_1 (Fig. (4b)). Riguardo alla definizione precedente, se il GE può fornire una rotazione di $\alpha_N = 45^\circ$ all’estremità libera in condizioni limite di snervamento, qualora sottoposto ad un momento flettente adeguato \mathbf{m} , allora $\alpha \in [\pm 45^\circ]$ coincide con lo spazio di lavoro W_s . Tale spazio di lavoro viene, in seguito, discretizzato in un numero finito N di posizioni angolari α_k (dove $k = 1, \dots, N$). Ovviamente, ad una data suddivisione di W_s corrisponde una ben definita discretizzazione dei carichi. Ad esempio, la posizione iniziale definita dal punto O_1 per $\alpha_1 = 0$, una qualsiasi posizione intermedia definita dal punto O_i per α_i e la posizione finale O_N per $\alpha_N = 45^\circ$ sono raggiunte qualora $\mathbf{m}_1 = 0$, $0 < \mathbf{m}_k < \mathbf{m}_N$ e $\mathbf{m}_N = \mathbf{m}$ rispettivamente. In pratica, si suppone che il GE sia sottoposto a grandi spostamenti principali ma a piccoli spostamenti secondari.

Valutazione di matrici di cedevolezza normalizzate

Una volta definita una discretizzazione dello spazio di lavoro, si propone di valutare una matrice di cedevolezza *locale* \mathbf{C}_k per ogni configurazione del giunto. Come detto, \mathbf{C}_k presenta elementi di dimensioni fisiche non uniformi. Pertanto, per potere considerare quantità adimensionali, si definisce un momento di inerzia equivalente:

$$I^* = l/EC_{\theta yMy} \quad (2)$$

dove $C_{\theta yMy}$ indica la cedevolezza principale del GE.

Similmente, come suggerito in [16], gli spostamenti e le lunghezze vengono normalizzati rispetto alla lunghezza totale del giunto l (si faccia riferimento a Fig. (2)), le forze rispetto a EI^*/l^2 , ed i momenti rispetto a EI^*/l , cosicchè la cedevolezza principale risulta $\tilde{C}_{\theta yMy} =$

1. Ogni matrice $\tilde{\mathbf{C}}_k$ è poi calcolata attraverso i seguenti passaggi concettuali:

- Si applica una frazione del carico principale massimo \mathbf{m} , in modo da raggiungere la k -esima configurazione, dove $k = 1, \dots, N$.
- Si applica una piccola variazione ad un carico secondario mantenendo il carico principale imposto precedentemente. Si misurano quindi gli spostamenti generati $\Delta \mathbf{u}$ (3 traslazioni e 3 rotazioni). Tali spostamenti sono utilizzati per il calcolo delle sei componenti di $\tilde{\mathbf{C}}_k$ relative alla particolare condizione di carico in esame (i.e. k -esima configurazione). Si definisce “piccola”, una variazione di carico che genera un piccolo spostamento [23].
- Si ripete la procedura per ogni componente di carico secondario, in modo da calcolare l'intera matrice di cedevolezza relativa alla k -esima configurazione.

La matrice di cedevolezza locale normalizzata è divisa in due sotto-matrici, $\tilde{\mathbf{C}}_{Rk}$ e $\tilde{\mathbf{C}}_{Tk}$, contenenti rispettivamente i coefficienti relativi a spostamenti angolari e a spostamenti lineari, lungo le direzioni di riferimento.

$$[\tilde{\mathbf{C}}]_k = \begin{bmatrix} \tilde{C}_{Rk} \\ \tilde{C}_{Tk} \end{bmatrix} \quad (3)$$

$$[\tilde{\mathbf{C}}]_{Rk} = \tilde{C}_{i,j}, \quad i = 1, \dots, 3; \quad j = 1, \dots, 6.$$

$$[\tilde{\mathbf{C}}]_{Tk} = \tilde{C}_{i,j}, \quad i = 3, \dots, 6; \quad j = 1, \dots, 6.$$

La stessa procedura viene utilizzata per le N suddivisioni dello spazio di lavoro. Si noti che: 1) è possibile ottenere una migliore accuratezza solo con un numero elevato N di configurazioni di giunto; 2) la scelta del sistema di riferimento rispetto al quale sono calcolate le \mathbf{C}_k deve essere congruente ed invariante durante l'intero processo di valutazione [22, 24]. Come esempio, la Fig. (5a) mostra la configurazione deformata (rotazione principale di 45°) del giunto *SPIR* se sottoposto unicamente a carico principale, mentre la Fig. (5b) mostra la configurazione deformata dello stesso giunto *SPIR* qualora soggetto a carico principale e a coppia (secondaria) lungo la direzione x (sistema di riferimento mostrato in Fig. (2c)).

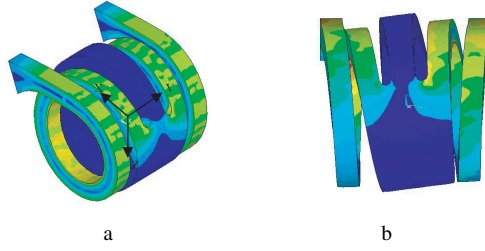


Figura 5. Esempio di FEA sul giunto *SPIR*. Configurazione deformate sotto l'applicazione di un carico (coppia) principale (a) oppure di un carico (coppia) principale + una coppia \mathbf{m}_x (b). Allo scopo di ridurre il costo computazionale, il modello FEA del giunto non include le aree in cui la deformazione risulta trascurabile.

Definizione e valutazione di indici di qualità locale e globale

Qualora si sia valutata la matrice di cedevolezza locale $\tilde{\mathbf{C}}_k$ per ognuna delle configurazioni in esame, si possono definire una serie di IQL utili nella valutazione di differenti GE. Gli indici proposti sono basati sulla norma matriciale di Frobenius, cioè:

$$\mathbf{A}_F = \sqrt{\frac{1}{n} \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2} \quad (4)$$

che, intuitivamente, indica l'ampiezza degli elementi di una generica matrice $\mathbf{A} \, m \times n$.

Nello specifico, si definiscono un IQL rotazionale I_r ed un IQL trasazionale I_t come:

$$I_{Rk} = \tilde{\mathbf{C}}_{\mathbf{R}k} \, F \quad I_{Tk} = \tilde{\mathbf{C}}_{\mathbf{T}k} \, F \quad (5)$$

Un minor IQL indica un migliore comportamento locale del GE.

La fase finale dell'approccio comparativo consiste nella definizione e nella valutazione di indici di qualità globale (IQG) che riassumano la qualità complessiva del giunto in termini di cedevolezza selettiva all'interno dello spazio di lavoro W_s :

$$I_{Rg} = \frac{\sum_{i=1}^N I_{Rk}}{N} \quad I_{Tg} = \frac{\sum_{i=1}^N I_{Tk}}{N} \quad (6)$$

Un minor IQG indica un migliore comportamento globale del GE.

Si noti che: 1) piccole rotazioni secondarie a livello del giunto possono essere notevolmente amplificate al termine di catene articolate seriali, pertanto la valutazione di I_{Rg} è di solito più significativa; 2) la definizione di indici globali è apprezzabile in termini di sintesi del metodo. Nonostante ciò, un'analisi dell'andamento degli IQL, all'interno dello spazio di lavoro, può far emergere effetti pericolosi (ad esempio effetti di *load stiffening*) non evidenziabili dalla sola analisi degli IQG.

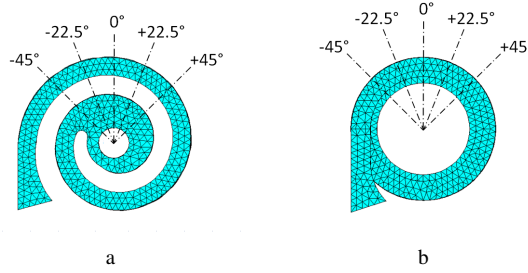


Figura 6. Discretizzazione dello spazio di lavoro: *SPIR* (a) e *HEL* (b) rispettivamente.

3. VALIDAZIONE DEL METODO

Per testare l'utilità della procedura proposta, si effettua una valutazione comparativa dei GE per applicazioni robotiche riportati in Fig. (2).

I giunti J_C e J_R sono semplici travi snelle a sezione circolare e rettangolare rispettivamente. In questo contesto, sono entrambe trattati come GE rotoidali (ad un singolo g.d.l.). Ovviamente, il comportamento di J_R è migliore in termini di cedevolezza selettiva se comparato a J_C , utilizzabile, infatti, come GE sferico [8].

I giunti *SPIR* e *HEL* sono stati progettati per essere utilizzati in dita robotiche articolate, avendo dimensioni comparabili con le articolazioni delle dita umane. Il giunto *SPIR* è basato su una coppia di molle planari con geometria a spirale, mentre il giunto *HEL* è basato su una coppia di molle elicoidali aventi l'asse dell'elica parallelo alla direzione principale di spostamento (asse *compliant*).

Il confronto è stato realizzato mediante FEA, utilizzando il software ANSYS Release 12.0. Il materiale utilizzato nelle simulazioni, denominato *Fullcure 720*, è utilizzato in stampanti 3D per la prototipazione rapida e presenta un modulo di Young $E = 2870\text{MPa}$ ed un coefficiente di Poisson $\nu = 0.33$. L'elemento scelto per questa analisi è *SOLID187*. Questo elemento a 10 nodi ben si adatta a modellare mesh irregolari. In riferimento alle tecniche di controllo della soluzione, è stata realizzata un'analisi statica. Lo spazio di lavoro dei GE è stato suddiviso in $N = 5$ configurazioni: -45° , -22.5° , 0° (configurazione indeformata), $+22.5^\circ$, $+45^\circ$ (il segno positivo si riferisce alla rotazione in senso orario, Fig. (6)).

Gli IQG ottenuti dall'analisi sono riportati in Tabella 1, mentre l'andamento degli IQL è riportato in Fig. (7).

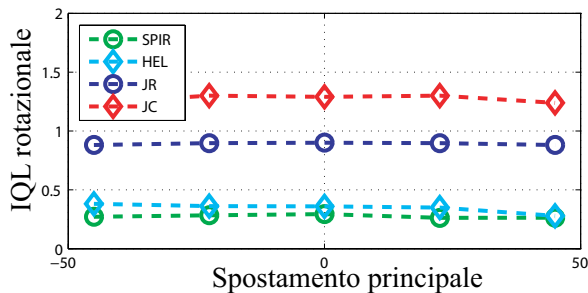
Con riferimento al confronto tra J_R e J_C , la Tabella 1 evidenzia valori più bassi degli IQG correlati a J_R . In aggiunta, la Fig. (7) mostra l'andamento degli IQL all'interno dello

Table 1: Indici globali per i giunti *SPIR* e *HEL*.

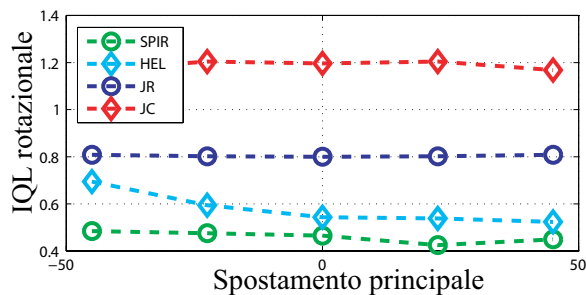
	I_{Tg}	I_{Rg}		I_{Tg}	I_{Rg}
J_C	1.2736	1.1880	<i>SPIR</i>	0.2749	0.4596
J_R	0.8911	0.8041	<i>HEL</i>	0.3467	0.5786

spazio di lavoro del giunto, confermando che J_C è caratterizzato da una cedevolezza selettiva più bassa. Con riferimento al confronto tra i quattro GE, i valori di tutti gli IQG mostrano prestazioni migliori per il giunto *SPIR*.

Questo risultato viene confermato dall'analisi dell'andamento degli IQL presentata in Fig. (7). L'analisi è stata limitata ad un intervallo di rotazione di $\pm 45^\circ$. Tuttavia, come mostrato in Fig. (8), i giunti *SPIR* e *HEL* sono in grado di ruotare di 90° in senso orario e di 45° in senso antiorario. Le tipologie di GE presentate in questa memoria sono state utilizzate per produrre il dito robotico mono-pezzo mostrato in Fig. (9).



a Andamento dell'IQL I_{Rk} per SPIR e HEL.



b Andamento dell'IQL I_{Tk} per SPIR e HEL.

Figura 7. Andamento degli indici di qualità locale (IQL) per i giunti *SPIR* e *HEL* (Fig. (2c) and (2d)).

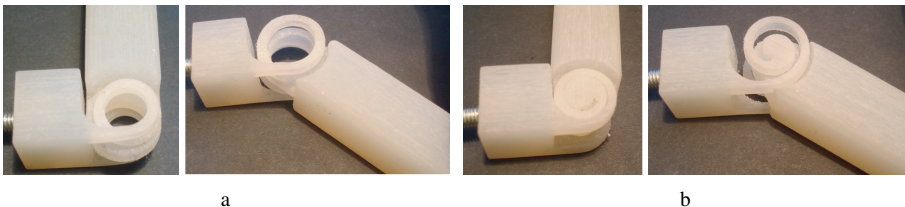


Figura 8. GE deformati: *SPIR* (a), *HEL* (b).

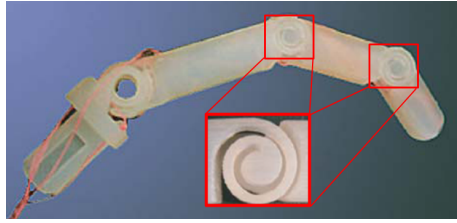


Figura 9. Dito robotico mono-pezzo. Utilizzo di giunti elastici piani a spirale.

4. DISCUSSIONE E CONCLUSIONI

I GE comunemente adottati nei svariati ambiti della tecnica sono in grado di fornire piccoli spostamenti. Per contro, molte strutture robotiche necessitano di spostamenti molto più ampi. Ad esempio, l'applicazione di GE in dita robotiche richiede l'utilizzo di morfologie di giunto capaci di rotazioni di circa 90° , dieci volte superiori rispetto alle rotazioni di cui sono passibili i comuni GE. Generalmente, giunti adatti a grandi spostamenti si possono ottenere incrementando l'estensione della regione di spazio occupata da materiale elastico sottoposto a deformazioni imposte (e.g. utilizzo di travi snelle ad elevata lunghezza). Tuttavia, tale soluzione incrementa inevitabilmente la suscettibilità a spostamenti secondari. Questo inconveniente rappresenta un fattore critico a sfavore dell'applicazione di GE in strutture robotiche articolate ove piccoli spostamenti secondari, a livello di giunto, possono essere enormemente amplificati al termine di una catena seriale (specialmente nel caso in cui siano presenti numerosi link di elevata lunghezza).

In pratica, le evidenti difficoltà nel progetto e nella valutazione del giunto, nonché le elevate sensibilità ai carichi secondari, porterebbero a sconsigliare l'utilizzo di GE in robotica.

Queste considerazioni possono, tuttavia, essere rilette in una prospettiva più ampia che includa i concetti di economicità globale di sistema e di ammissibilità di errori più o meno rilevanti. Infatti, esistono casi in cui la presenza di spostamenti indesiderati può essere limitata ed, in ogni caso, non compromette la capacità di svolgere con successo i compiti preposti. Considerando, per esempio, mani robotiche e protesiche con dita articolate, durante la traiettoria di avvicinamento all'oggetto, le dita stesse si comportano come catene seriali indipendenti e soggette a carichi molto limitati; in questa fase, gli errori di traiettoria dovuti a spostamenti secondari risultano limitati ed accettabili. A seguito della presa e dell'applicazione dei contatti tra mano e oggetto, il sistema oggetto-dita può essere concepito come una struttura parallela [25] nella quale le dita contribuiscono alla stabilità ed alla robustezza della presa, mitigando gli effetti degli spostamenti secondari dovuti al singolo giunto.

Queste osservazioni qualitative sottolineano l'importanza, nella pratica, dei metodi di confronto presentati in questa memoria ed atti a stabilire quale GE meglio si adatti ai requisiti di una data applicazione. In conclusione, il criterio proposto è basato sulla valutazione di indici di qualità locali e globali che quantificano la cedevolezza selettiva del giunto. L'approccio è stato validato comparando due GE la cui soluzione analitica è nota da letteratura e due GE di morfologia non banale. In particolare, i giunti *SPIR* e *HEL* sono stati ideati per utilizzo su dita robotiche, ove l'inconveniente degli effetti parassiti può essere accettato a fronte di una minore complessità costruttiva.

Ringraziamenti

Questa ricerca è stata finanziata da *EC Seventh Framework Programme* (FP7) all'interno del progetto *DEXMART* (progetto n. 216239) e dal MIUR all'interno del PRIN2007CCRNFA-004 (*SICURA project*). Si ringraziano: Prof. Claudio Melchiorri, Ing. Gianluca Palli e Ing. Gianni Borghesan.

BIBLIOGRAFIA

- [1] Ananthasuresh, G., and Kota, S., 1995. "Designing compliant mechanisms.". *Mech. Eng.*, **117**, p. 936.
- [2] Berselli, G., Piccinini, M., and Vassura, G., 2011. "Comparative evaluation of the selective compliance in elastic joints for robotic structures". In Proc. IEEE Int. Conf. on Robotics and Automation, ICRA.
- [3] Trease, B. P., Moon, Y.-M., and Kota, S., 2005. "Design of large-displacement compliant joints". *Journal of Mechanical Design*, **127**(4), pp. 788–798.
- [4] Lotti, F., and Vassura, G., 2002. "A novel approach to mechanical design of articulated fingers for robotic hands". In Intelligent Robots and Systems, 2002. (IROS 2002). 2002 IEEE/RSJ International Conference on.
- [5] Carrozza, M. C., Cappiello, G., Stellan, G., Zaccone, F., Vecchi, F., Micera, S., and Dario, P., 2005. "A cosmetic prosthetic hand with tendon driven under-actuated mechanism and compliant joints: Ongoing research and preliminary results". In ICRA, pp. 2661–2666.
- [6] Hasse, A., and Campanile, L., 2009. "Design of compliant mechanisms with selective compliance.". *Smart Mater. Struct.*, **18**, p. 115016 (10pp).
- [7] Hoover, A., and Fearing, R., 2009. "Analysis of off-axis performance of compliant mechanisms with applications to mobile millirobot design". pp. 2770–2776.
- [8] Lobontiu, N., 2003. *Compliant Mechanisms: Design of flexures hinges*. CRC press.
- [9] Howell, L., 2001. *Compliant Mechanisms*. Jhon Wiley and Sons.
- [10] Goldfarb, M., and Speich, J. E., 1999. "A well-behaved revolute flexure joint for compliant mechanism design". *Journal of Mechanical Design*, **121**(3), pp. 424–429.
- [11] Hatanaka, M., and Cutkosky, M., 2003. "Process planning for embedding flexible materials in multi-material prototypes". In Proceedings of DETC06, ASME Design Engineering Technical Conferences, Philadelphia.
- [12] Dollar, A. M., and Howe, R. D., 2006. "A robust compliant grasper via shape deposition manufacturing". *IEEE/ASME Transactions on Mechatronics*, **11**(2).
- [13] De Bona, F., and Munteanu, M., 2005. "Optimized flexural hinges for compliant micromechanisms.". *Analog integrated circuits and signal processing*, **44**, pp. 163–174.
- [14] Vallance, R. R., Haghighian, B., and Marsh, E. R., 2008. "A unified geometric model for designing elastic pivots". *Precision Engineering*, **32**(4), pp. 278 – 288.
- [15] Zhou, H., and Ting, K.-L., 2006. "Shape and size synthesis of compliant mechanisms using wide curve theory". *Journal of Mechanical Design*, **128**(3), pp. 551–558.
- [16] Awtar, S., Slocum, A., and Sevincer, E., 2006. "Characteristics of beam-based flexure modules.". *ASME Journal of Mechanical Design*, **129** (6), pp. 625–639.
- [17] Mattiasson, K., 1981. "Numerical results from large deflection beam and frame problems analyzed by means of elliptic integrals.". *Int. J. Numer. Methods Eng.*, **17**, p. 145153.

- [18] Zelenika, S., and DeBona, F., 2002. "Analytical and experimental characterization of high precision flexural pivots subjected to lateral loads.". *Precis. Eng.*, **26**, p. 381388.
- [19] Patterson, T., and Lipkin, H., 1993. "Structure of robot compliance". *Journal of Mechanical Design*, **115**(3), pp. 576–580.
- [20] Lipkin, H., and Patterson, T., 1992. "Geometrical properties of modelled robot elasticity: part i - decomposition". *Proceedings of ASME, Design Technical Conference and Computers in Engineering Conference*, **45**, pp. 179–185.
- [21] Lobontiu, N., and Paine, J. S. N., 2002. "Design of circular cross-section corner-filletted flexure hinges for three-dimensional compliant mechanisms". *Journal of Mechanical Design*, **124**(3), pp. 479–484.
- [22] Howard, S., Zefran, M., and Kumar, V., 1998. "On the 6 × 6 cartesian stiffness matrix for three-dimensional motions". *Mechanism and Machine Theory*, **33**(4), pp. 389 – 408.
- [23] Kvecses, J., and Angeles, J., 2007. "The stiffness matrix in elastically articulated rigid-body systems". *Multibody System Dynamics*, **18**(2), pp. 169–184.
- [24] Zhang, S., and Fasse, E. D., 2001. "A finite-element-based method to determine the spatial stiffness properties of a notch hinge". *Journal of Mechanical Design*, **123**(1), pp. 141–147.
- [25] Murray, R., Li, Z., and Sastry, S., 1994. *A Mathematical Introduction to Robotic Manipulation*. CRC Press.

UNDER-CONSTRAINED CABLE-DRIVEN PARALLEL ROBOTS

Marco Carricato

DIEM - Dept. of Mechanical, Nuclear, Aviation and Metallurgical Engineering,
University of Bologna, Italy
E-mail: marco.carricato@unibo.it

Abstract. *This paper studies cable-driven parallel robots with less than six cables, in crane configuration. A major challenge in the study of these manipulators is the intrinsic coupling between kinematics and statics, which must be dealt with simultaneously. A geometrico-static model is provided, together with a general procedure aimed at effectively solving the inverse and direct position problems. The stability of equilibrium is assessed within the framework of a constrained optimization problem, for which a purely algebraic formulation is provided. A spatial robot with three cables is studied as an application example.*

Keywords: *Cable-driven parallel robots, underactuated robots, kinematic analysis, static analysis, stability analysis.*

1. INTRODUCTION

Cable-driven parallel robots (CDPRs) strengthen classic advantages characterizing closed-chain architectures versus serial ones, like reduced masses and inertias, a larger payload to robot weight ratio, high dynamic performances, improved motor efficiency, etc., while providing peculiar advantages, such as a larger workspace, reduced manufacturing and maintenance costs, ease of assembly and disassembly, high transportability, superior modularity and reconfigurability. CDPRs are ordinarily referred to as *fully-constrained* or *under-constrained*, depending on whether all six degrees of freedom (dofs) of the moving platform are controlled or not [1]. It is well known that, in the general case, fully-constrained CDPRs require at least seven cables, and only six in crane configuration, namely when gravity acts as an additional cable. The employ of under-constrained CDPRs with less than six cables is justified in several applications (such as, for instance, rescue, service or rehabilitation operations), in which the task to be performed requires a limited number of controlled freedoms (only n dofs of the end-effector may be governed by n cables) or a limitation of dexterity is acceptable in order to decrease complexity, cost, set-up time, likelihood of cable interference, etc. Furthermore, a theoretically fully-constrained CDPR may operate, in appreciable parts of its geometric workspace, as an under-constrained robot, namely when a full restraint of the end-effector may not be achieved because it would require a negative tension in one or more cables.

Even though the above considerations motivate a careful study of under-constrained CD-

PRs, little research has been dedicated to them [2–7], whereas a much richer literature exists concerning fully-constrained robots [1, 8–35]. A major challenge in the study of under-constrained CDPRs consists in the intrinsic coupling between kinematics and statics (or dynamics). When a fully-constrained CDPR operates in the portion of its workspace in which the required set of output wrenches is guaranteed to be applicable with purely tensile cable forces, the posture of the end-effector is determined, in a purely geometrical way, by assigning cable lengths. Conversely, for an under-constrained CDPR, when the actuators are locked and the cable lengths are assigned, the end-effector is still movable, so that the actual configuration is determined by the applied forces. As a consequence, the end-effector posture depends on both cable lengths and equilibrium equations. The necessity to deal with kinematics and statics simultaneously increases the complexity of position problems, which are aimed at determining the overall robot configuration when a set of n variables is assigned. The solution of these problems is significantly more difficult than analogous tasks concerning rigid-link parallel manipulators. Moreover, as the end-effector pose depends on the applied load, it may change due to external disturbances, so that it is fundamental to investigate equilibrium stability. An equilibrium configuration is, actually, *feasible* only if both cable tensions are positive and the configuration is stable.

This paper presents a preliminary study of the kinematics and statics of under-constrained nn -CDPRs, namely parallel robots in which a fixed base and a mobile platform are connected to each other by n cables, with $n \leq 5$ and the anchor points on the base and the platform being generally distinct. A geometrico-static model is presented, together with a general procedure aimed at effectively solving the inverse and direct position problems. These consist in determining the overall robot configuration and cable tensions when, respectively, a set of n platform posture coordinates or the n cable lengths are assigned. The problem of equilibrium stability is formulated as a constrained optimization problem, and a purely algebraic method is provided that rules out the need of differentiation. The geometrico-static study of a general 33-CDPR is outlined as an application example.

2. GEOMETRICO-STATIC MODEL

Let a mobile platform be connected to a fixed base by n cables, with $2 \leq n \leq 5$. A_i and B_i are, respectively, the exit point on the base and the anchor point on the platform, and $\mathbf{s}_i = B_i - A_i$ (Fig. 1). The set \mathcal{C} of geometrical constraints imposed on the platform comprises the relations

$$|\mathbf{s}_i| = \sqrt{\mathbf{s}_i \cdot \mathbf{s}_i} = \rho_i, \quad i = 1 \dots n, \quad (1)$$

where ρ_i is the length of the i th-cable, which is assumed, for apparent practical reasons, strictly positive (as a consequence, $B_i \neq A_i$).

Since only n geometrical constraints are enforced, the platform preserves $6 - n$ degrees of freedom, with its posture being determined by equilibrium laws. If $Q\$\mathbf{e}$, with $Q > 0$, is an arbitrary external wrench acting on the platform (including inertia forces, in case of dynamic conditions) and $(\tau_i/\rho_i)\$\mathbf{i}$ is the force exerted by the i th cable ($\$\mathbf{e}$ and $\$\mathbf{i}/\rho_i$ are assumed to be unit screws), then

$$\sum_{i=1}^n \frac{\tau_i}{\rho_i} \$\mathbf{i} + Q\$\mathbf{e} = \mathbf{0}, \quad (2)$$

with

$$\tau_i \geq 0, \quad i = 1 \dots n. \quad (3)$$

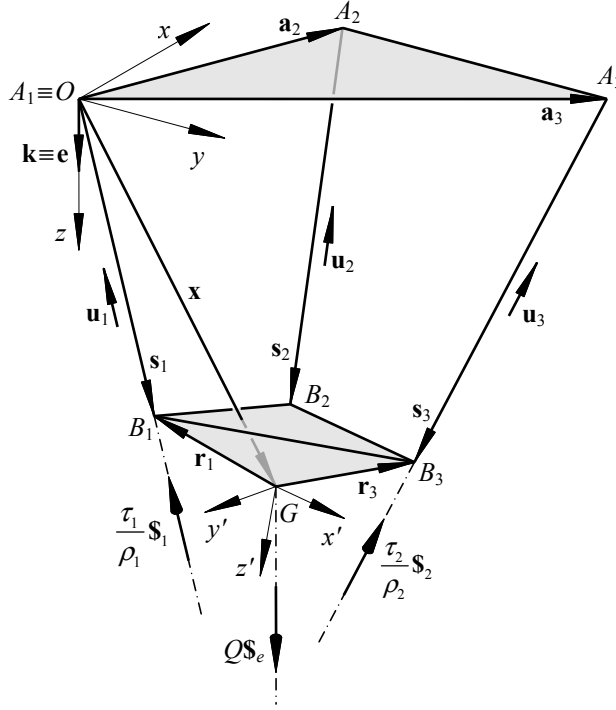


Figure 1: Geometric model of a cable-driven parallel robot with three cables.

Equations (1)-(2) amount to $6 + n$ scalar relations involving $6 + 2n$ variables, namely, the cable tensions and lengths, and the variables used to parameterize the platform posture. In general, a finite set of system configurations may be determined if n of such variables are assigned. In this paper, only *static* equilibrium is considered and $Q\mathbf{s}_e$ is assumed to be a *constant* force applied on a point G of the platform (e.g., the platform weight acting through its center of mass). Hence, Eqs. (1)-(2) are algebraic, or may be easily rendered so.

If Eq. (2) is written as

$$\underbrace{\begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_n & \mathbf{s}_e \end{bmatrix}}_{\mathbf{M}} \begin{bmatrix} (\tau_1/\rho_1) \\ \vdots \\ (\tau_n/\rho_n) \\ Q \end{bmatrix} = \mathbf{0}, \quad (4)$$

\mathbf{M} is a $6 \times (n+1)$ matrix only depending on the platform posture and equilibrium is possible only if

$$\text{rank}(\mathbf{M}) \leq n, \quad (5)$$

namely, if the cables and the line of action of \mathbf{s}_e span the same n -dimensional system of lines. Within the domain of rigid-body mechanics, the problem is *statically determinate* if the equality holds, *indeterminate* otherwise. In the former case, it is always possible to replace Eq. (4) with $6 - n$ scalar relations that do not contain the unknowns $\tau_i, i = 1 \dots n$. In

fact, the linear dependence of $\$1, \dots, \n and $\$e$ is a purely geometrical condition. A most straightforward strategy consists in computing cable tensions by way of n linearly independent relations chosen within Eq. (4), and then substituting them back into the remaining ones. A more convenient strategy consists in setting all $(n+1) \times (n+1)$ minors of \mathbf{M} equal to zero, which amounts to $\binom{6}{n+1}$ scalar relations, among which $6-n$ independent ones may be suitably chosen (the most appropriate choice depends on the specific problem at hand). Such an approach provides a wider set of linearly independent equations and it offers a better insight into their structure. Furthermore, since these relations do *not* comprise cable lengths, a partial decoupling of the system equations is achieved, with cable lengths only appearing in Eq. (1). Such an approach may also be applied when the problem is statically indeterminate.

Depending on the variables designated as input, one may tackle an inverse geometrico-static problem (IGP), if n variables concerning the platform posture are assigned, or a direct one (DGP), if cable lengths are given. The IGP takes particular advantage of the partial decoupling of the system equations, since, in this case, the $6-n$ configuration variables that are needed to fully determine the platform pose may be directly computed by way of (a minimum of $6-n$) relations emerging from Eq. (5). Subsequently, cable lengths may be directly computed by Eq. (1) and cable tensions may be obtained by a suitable set of linear independent relations chosen within Eq. (4). The set of *feasible* configurations consists of those for which cable tensions are non-negative (cf. Eq. (3)) and the platform equilibrium is stable (cf. §3.). The DGP poses remarkably more complex mathematical problems, since in this case the platform configuration must be determined by simultaneously solving both the relations emerging from Eq. (5) and the n relations in Eq. (1).

It must be said that Eq. (1) represents a set of *theoretical* constraints, since the *actual* constraint imposed by a generic cable consists in that

$$|\mathbf{s}_i| = \sqrt{\mathbf{s}_i \cdot \mathbf{s}_i} \leq \rho_i. \quad (6)$$

The above refinement causes no concern when the IGP is dealt with, for in this case the *theoretical* values of cable lengths are conveniently computed by Eq. (1), after the platform posture has been established. Conversely, when the DGP is tackled, cable lengths are assigned as inputs, and, *a priori*, nothing assures that *all* cables are called upon to sustain the load. Indeed, if a subset \mathcal{W} of cable indexes exists such that $\text{card}(\mathcal{W}) < n$ and $\$e \in \text{span}(\$j, j \in \mathcal{W})$, equilibrium configurations possibly exist such that $|\mathbf{s}_k| < \rho_k$ and thus $\tau_k = 0$, for all $k \notin \mathcal{W}$. These are legitimate solutions of the problem at hand. It follows that the overall solution set is obtained by solving the DGP for *all possible* constraint sets $\{|\mathbf{s}_j| = \rho_j, j \in \mathcal{W}\}$, with $\mathcal{W} \subseteq \{1 \dots n\}$ and $\text{card}(\mathcal{W}) \leq n$, and by retaining, for each corresponding solution set, the solutions for which $|\mathbf{s}_k| < \rho_k, k \notin \mathcal{W}$, and Eq. (3) is verified. In general, this amounts to solving $\sum_{h=0}^{n-1} \binom{n}{n-h}$ DGPs.

A caveat is worth being mentioned. Equation (5) is a necessary, but *not* sufficient, condition for equilibrium. In very special conditions, it may happen that equilibrium is not possible, in spite of Eq. (5) being fulfilled and irrespective of the sign of cable tensions. In particular, this occurs if \mathbf{M} loses its full rank because a subset of its n first columns becomes linearly dependent, i.e. if the rank loss is ‘concentrated’ among the set of screws associated with the cable lines. In this situation, the rank of the block¹ $\mathbf{M}_{1\dots 6, 1\dots n}$ is, at the most, equal to $n-1$ and Eq. (2) may be satisfied only if $\text{rank}(\mathbf{M}) \leq n-1$, since, at the equilibrium,

¹The notation $\mathbf{M}_{hij,klm}$ denotes the block matrix obtained from rows h, i and j , and columns k, l and m , of \mathbf{M} . When all columns of \mathbf{M} are used, the corresponding subscripts are omitted.

the external wrench must, in any case, belong to the subspace generated by the cable lines. Events like these must be separately studied and dealt with. For this reason, a check of the rank of $\mathbf{M}_{1\dots 6,1\dots n}$ is advisable before attempting to solve for cable tensions.

Throughout the text, the following notation is adopted (Fig. 1). $Oxyz$ is a Cartesian coordinate frame fixed to the base, with \mathbf{i} , \mathbf{j} and \mathbf{k} being unit vectors along the coordinate axes. $Gx'y'z'$ is a Cartesian frame attached to the platform. \mathbf{e} is a unit vector directed as $\$e$, $\mathbf{x} = G - O$, $\mathbf{a}_i = A_i - O$, $\mathbf{r}_i = B_i - G$, $\mathbf{s}_i = B_i - A_i = \mathbf{x} + \mathbf{r}_i - \mathbf{a}_i$, $\mathbf{u}_i = (A_i - B_i)/\rho_i = -\mathbf{s}_i/\rho_i$ and $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, with $i, j = 1 \dots n$, $i \neq j$. Without loss of generality, O is chosen to coincide with A_1 (so that $\mathbf{a}_1 = \mathbf{0}$) and $\mathbf{k} = \mathbf{e}$. Furthermore, if \mathbf{b}_i is the projection of $B_i - G$ on $Gx'y'z'$, Φ is the array grouping the variables parameterizing the platform orientation with respect to the fixed frame, and $\mathbf{R}(\Phi)$ is the corresponding rotation matrix, then $\mathbf{r}_i = \mathbf{R}(\Phi) \mathbf{b}_i$. Finally, the platform posture is described by the array $\mathbf{X} = [\mathbf{x}; \Phi]$, with the components of \mathbf{x} in $Oxyz$ being denoted, for the sake of brevity, as x , y and z . If O is chosen as the reduction pole of moments, $\$i$ and $\$e$ may be respectively expressed, in axis coordinates, as $\$i = -[\mathbf{s}_i; \mathbf{a}_i \times \mathbf{s}_i]$ and $\$e = [\mathbf{e}; \mathbf{x} \times \mathbf{e}]$. Accordingly, \mathbf{M} becomes

$$\mathbf{M} = \begin{bmatrix} -(\mathbf{x} + \mathbf{r}_1) & \cdots & \mathbf{a}_i - (\mathbf{x} + \mathbf{r}_i) & \cdots & \mathbf{e} \\ \mathbf{0} & \cdots & -\mathbf{a}_i \times (\mathbf{x} + \mathbf{r}_i) & \cdots & \mathbf{x} \times \mathbf{e} \end{bmatrix}, \quad (7)$$

or, equivalently, by elementary column transformations,

$$\mathbf{M}' = \begin{bmatrix} \mathbf{x} + \mathbf{r}_1 & \cdots & \mathbf{r}_{i1} - \mathbf{a}_i & \cdots & \mathbf{e} \\ \mathbf{0} & \cdots & \mathbf{a}_i \times (\mathbf{x} + \mathbf{r}_i) & \cdots & \mathbf{x} \times \mathbf{e} \end{bmatrix}. \quad (8)$$

3. STABILITY OF EQUILIBRIUM

Let an equilibrium configuration $(\bar{\mathbf{X}}, \bar{\rho}_1 \dots \bar{\rho}_m)$ be considered, with m being the number of *active* constraints (i.e. the number of cables contributing to supporting the platform). By a convenient reordering of indexes, taut cables may be assumed to be the first m , with $m \leq n$. Since the platform preserves $6 - m$ dofs, it may displace under the effect of a change in the external force acting on it, while cable lengths remain unvaried (for the sake of simplicity, it is assumed that the number of cables in tension does not change because of the perturbation, which is reasonable, but not necessarily true). The problem of assessing equilibrium stability is in order. In particular, G may generally move within a closed region of \mathbb{R}^3 (in some cases, a surface or a curve).

If g is the frontier of this region, the equilibrium is stable any time the potential energy U associated with the external force $Q\$e$, namely $-Q\mathbf{e} \cdot \mathbf{x}$, is at a minimum on g . Loosely speaking, the platform is at rest in all points \bar{G} of g in which the variety tangent to g is perpendicular to \mathbf{e} , with the equilibrium being stable if and only if a neighborhood $W_{\bar{G}}$ of \bar{G} exists such that $(P - \bar{G}) \cdot \mathbf{e} < 0$, for all $P \in (g \cap W_{\bar{G}})$. In such a condition, when the platform displaces under the effect of a perturbation, the original configuration is restored if the perturbation ceases. Figure 2 helps to depict this concept. The figure shows the locus g of the positions that G may assume for an exemplifying 22-CDPR, under the constraints (1) and with $m = n = 2$. If the platform is thought of as the coupler of a four-bar linkage whose grounded links are the cables (with assigned lengths), g is the coupler curve of G , namely a bicursal sextic. The stationary configurations of G are the points of g in which the tangent line is perpendicular to \mathbf{e} , with U being at a minimum in \bar{G}_2 and \bar{G}_4 . These are the stable equilibrium poses. Of course, since cable tensions are negative in the configurations lying

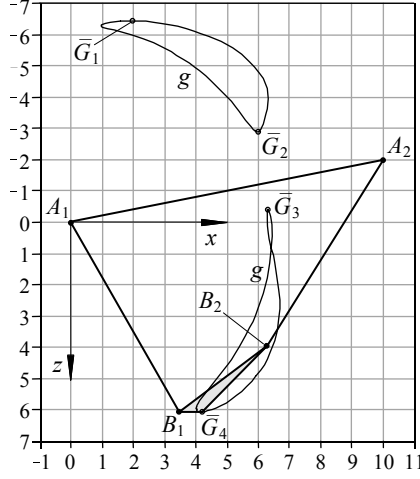


Figure 2: Equilibrium configurations of a 22-CDPR with:
 $\mathbf{a}_2 = [10; 0; -2]$, $\mathbf{b}_1 = [-0.5; 0; -0.5]$, $\mathbf{b}_2 = [3; 0; 0]$, and $\rho_1 = \rho_2 = 7$.

above the base and positive otherwise, \bar{G}_4 is *de facto* the only feasible configuration for the example at hand.

Finding the minima of a constrained function is a classical issue in optimization theory. An efficient algorithmic formalization is presented in the following. At equilibrium, the variation of the potential energy of the platform due to a virtual displacement of it must be zero. Such a variation is the opposite of the work carried out by all forces acting on the platform, namely

$$\delta L = - \sum_{i=1}^m \tau_i \mathbf{u}_i \cdot \delta B_i - Q \mathbf{e} \cdot \delta G = 0. \quad (9)$$

If $\delta \mathbf{x}$ and $\delta \Theta$ are, respectively, the virtual displacement of G and the virtual rotation of the platform, then $\delta G = \delta \mathbf{x}$ and $\delta B_i = \delta \mathbf{s}_i = \delta \mathbf{x} + \delta \Theta \times \mathbf{r}_i$, and thus

$$\delta L = - \left(\sum_{i=1}^m \tau_i \mathbf{u}_i + Q \mathbf{e} \right) \cdot \delta \mathbf{x} - \left(\sum_{i=1}^m \tau_i \mathbf{r}_i \times \mathbf{u}_i \right) \cdot \delta \Theta = \mathbf{f} \cdot \delta \mathbf{x} + \mathbf{m} \cdot \delta \Theta = 0. \quad (10)$$

Equation (10), from which \mathbf{f} and \mathbf{m} are inferred to be zero, is clearly equivalent to Eq. (2), by letting $n = m$. Since, for $\rho_i = \bar{\rho}_i$, $\delta(|\mathbf{s}_i| - \rho_i) = \delta|\mathbf{s}_i|$ and

$$\delta|\mathbf{s}_i| = \frac{\mathbf{s}_i \cdot \delta \mathbf{s}_i}{\rho_i} = \frac{\mathbf{s}_i \cdot \delta \mathbf{x} + \mathbf{r}_i \times \mathbf{s}_i \cdot \delta \Theta}{\rho_i} = -(\mathbf{u}_i \cdot \delta \mathbf{x} + \mathbf{r}_i \times \mathbf{u}_i \cdot \delta \Theta), \quad (11)$$

δL may be written as

$$\delta L = -Q \mathbf{e} \cdot \delta \mathbf{x} + \sum_{i=1}^m \tau_i \delta(|\mathbf{s}_i| - \rho_i), \quad (12)$$

i.e., as the virtual variation of the Lagrange function

$$L = -Q\mathbf{e} \cdot \mathbf{x} + \sum_{i=1}^m \tau_i (|\mathbf{s}_i| - \rho_i). \quad (13)$$

The Lagrange multipliers in L coincide with the cable tensions, namely, with the forces necessary to impose the geometrical constraints [36].² The above observation is useful, since it allows the stability characteristics of the equilibrium to be assessed by evaluating the definiteness of the reduced Hessian \mathbf{H}_r of L , namely, the Hessian of L taken with respect to the configuration variables, further projected on the tangent space of the constraints \mathcal{C} [37]. An algebraic expression of \mathbf{H}_r is derived hereafter.

The second-order variation of δL is given by

$$\delta^2 L = -Q\mathbf{e} \cdot \delta^2 \mathbf{x} + \sum_{i=1}^m \tau_i \frac{\delta \mathbf{s}_i \cdot \delta \mathbf{s}_i}{\rho_i} + \sum_{i=1}^m \tau_i \frac{\mathbf{s}_i \cdot \delta^2 \mathbf{s}_i}{\rho_i}, \quad (14)$$

with

$$\delta^2 \mathbf{s}_i = \delta^2 \mathbf{x} + \delta^2 \boldsymbol{\Theta} \times \mathbf{r}_i + \delta \boldsymbol{\Theta} \times (\delta \boldsymbol{\Theta} \times \mathbf{r}_i). \quad (15)$$

Substituting Eq. (15) in Eq. (14), and enforcing $\mathbf{f} = \mathbf{m} = \mathbf{0}$, yields

$$\delta^2 L = \sum_{i=1}^m \frac{\tau_i}{\rho_i} \{ \delta \mathbf{x} \cdot \delta \mathbf{x} - 2\delta \mathbf{x} \cdot (\mathbf{r}_i \times \delta \boldsymbol{\Theta}) - (\mathbf{r}_i \times \delta \boldsymbol{\Theta}) \cdot [(\mathbf{x} - \mathbf{a}_i) \times \delta \boldsymbol{\Theta}] \} \quad (16)$$

or, in matrix notation,

$$\delta^2 L = \sum_{i=1}^m \frac{\tau_i}{\rho_i} \left[\delta \mathbf{x}^T \delta \mathbf{x} - 2\delta \mathbf{x}^T \tilde{\mathbf{r}}_i \delta \boldsymbol{\Theta} + \delta \boldsymbol{\Theta}^T \tilde{\mathbf{r}}_i (\tilde{\mathbf{x}} - \tilde{\mathbf{a}}_i) \delta \boldsymbol{\Theta} \right], \quad (17)$$

where $\tilde{\mathbf{n}}$ denotes, for a generic vector \mathbf{n} , the skew-symmetric matrix expressing the operator $\mathbf{n} \times$. $\delta^2 L$ is a bilinear form in the twist space of the platform. If the platform virtual displacement is expressed, in ray coordinates, as $\delta \mathbf{t} = [\delta \mathbf{x}; \delta \boldsymbol{\Theta}]$, and \mathbf{I}_3 denotes the 3×3 identity matrix, the symmetric matrix associated with this form is

$$\mathbf{H}_p = \sum_{i=1}^m \frac{\tau_i}{\rho_i} \begin{bmatrix} \mathbf{I}_3 & -\tilde{\mathbf{r}}_i \\ \tilde{\mathbf{r}}_i & \frac{1}{2} [\tilde{\mathbf{r}}_i (\tilde{\mathbf{x}} - \tilde{\mathbf{a}}_i) + (\tilde{\mathbf{x}} - \tilde{\mathbf{a}}_i) \tilde{\mathbf{r}}_i] \end{bmatrix}, \quad (18)$$

which represents the pseudo-Hessian of L (\mathbf{H}_p is not a true and proper Hessian, since $\delta \boldsymbol{\Theta}$ is not generally integrable).

The tangent space of \mathcal{C} is obtained by setting Eq. (11) equal to zero for all values of i . In matrix notation, this amounts to

$$\mathbf{J}_p \delta \mathbf{t} = \begin{bmatrix} \mathbf{s}_1^T & (\mathbf{r}_1 \times \mathbf{s}_1)^T \\ \vdots & \vdots \\ \mathbf{s}_m^T & (\mathbf{r}_m \times \mathbf{s}_m)^T \end{bmatrix} \begin{bmatrix} \delta \mathbf{x} \\ \delta \boldsymbol{\Theta} \end{bmatrix} = \mathbf{0}, \quad (19)$$

²Notice that Eq. (9) plus the relations $\{\tau_i > 0, |\mathbf{s}_i| = \rho_i\}$ for $i = 1 \dots m$ and $\{\tau_i = 0, |\mathbf{s}_i| < \rho_i\}$ for $i = m+1 \dots n$ are equivalent to the Karush-Kuhn-Tucker conditions for the minimization of L under the constraints (6), provided that $\mathbf{s}_1, \dots, \mathbf{s}_m$ are linearly independent.

where the i th row of \mathbf{J}_p coincides with $-\$i$, expressed in axis coordinates and assuming G as the moment pole. \mathbf{J}_p is the pseudo-Jacobian of the constraint equations.

If \mathbf{N}_p is any $6 \times (6 - m)$ matrix whose columns generate the null space of \mathbf{J}_p , the reduced Hessian of \mathcal{C} is the following $(6 - m) \times (6 - m)$ matrix:

$$\mathbf{H}_r = \mathbf{N}_p^T \mathbf{H}_p \mathbf{N}_p. \quad (20)$$

A sufficient condition for the equilibrium to be stable consists in \mathbf{H}_r being positive definite.

It is worth remarking that, if \mathbf{E}_3 is the matrix such that $\delta\Theta = \mathbf{E}_3\delta\Phi$, the Hessian and the Jacobian derived by differentiating L and \mathcal{C} with respect to \mathbf{x} and Φ are equal, respectively, to $\mathbf{H} = \mathbf{E}_6^T \mathbf{H}_p \mathbf{E}_6$ and $\mathbf{J} = \mathbf{J}_p \mathbf{E}_6$, where $\mathbf{E}_6 = [\mathbf{I}_3, \mathbf{0}_3; \mathbf{0}_3, \mathbf{E}_3]$ and $\mathbf{0}_3$ is the 3×3 zero matrix. If the null space of \mathbf{J} is generated by $\mathbf{N} = \mathbf{E}_6^{-1} \mathbf{N}_p$, then $\mathbf{H}_r = \mathbf{N}_p^T \mathbf{H}_p \mathbf{N}_p = \mathbf{N}^T \mathbf{H} \mathbf{N}$.

The method proposed in [7] differs from the one presented here in that it determines the stability of equilibrium by looking at the Hessian of an *unconstrained* potential, explicitly expressed as a function of m independent coordinates. Such a mapping, however, is generally very difficult to obtain (indeed, Michael *et al.* [7] apply important simplifications on the geometry of the robot) and extensive differential symbolic computation is needed. The advantage of the method described here consists in that \mathbf{H}_r is computed in a purely algebraic way, with no need to perform any differentiation, and it may be very simply applied to the most general cases.

4. APPLICATION EXAMPLE: THE 33-CDPR

Due to space limitations, only a brief outline of the IGP and the DGP of the 33-CDPR is sketched hereafter. Technical details and convenient discussions will be provided in future papers, as well as the study of the cases $n = 2, 4, 5$.

When $n = 3$, Eq. (5) is satisfied and $\text{rank}(\mathbf{M}) = 3$ only if $\$1$, $\$2$, $\$3$ and $\$e$ belong to the same tridimensional subspace of lines [38]. Letting all 4×4 minors of \mathbf{M}' vanish leads to 15 polynomial equations in \mathbf{x} and Φ of the form $p_j = 0$. We look for the variety V of the ideal generated by such equations. If three configuration variables are known (as in the IGP), any three p_j , say p_l, p_h, p_k , may be chosen and a corresponding (generally zero-dimensional) variety V_{lhk} is obtained. V is the intersection of the five varieties that may be generated this way. Clearly, a primary objective of the solving algorithm consists in limiting the number of varieties to be computed to the lowest possible value, possibly to just one.

It is useful observing that letting $B_i \equiv A_i$ causes the i th column of \mathbf{M} to vanish (since $\mathbf{s}_i = \mathbf{a}_i \times \mathbf{s}_i = \mathbf{0}$) and, thus, it causes all 4×4 minors of \mathbf{M} (and thus of \mathbf{M}') to be zero. It follows that, when a configuration for which $B_i \equiv A_i$ is compatible with the assigned constraints, it necessarily belongs to V : we call it a *trivial* solution and we need to discard it (cf. §2.). This observation is particularly important for the IGP with assigned orientation. In this case, in fact, it is always possible to displace the platform (with a given orientation) so as to superimpose B_i onto A_i . Consequently, all varieties V_{lhk} necessarily contain the trivial solutions corresponding to $B_i \equiv A_i$, namely $\bar{\mathbf{x}}_i = \mathbf{a}_i - \mathbf{r}_i, i = 1 \dots 3$.

Inverse geometrico-static problem

When the orientation is assigned, all vectors \mathbf{r}_i , $i = 1 \dots 3$, are known. If the equations

$$p_1 := \det \mathbf{M}'_{1236} = 0, \quad (21a)$$

$$p_2 := \det \mathbf{M}'_{1235} = 0, \quad (21b)$$

$$p_3 := \det \mathbf{M}'_{1234} = 0, \quad (21c)$$

are considered, it may be proven that $V \equiv V_{123}$. Such equations comprise the lowest-degree polynomials among all minors of \mathbf{M}' . In particular, p_1 is quadratic in x and y , whereas p_2 and p_3 are quadratic in x , y and z . By eliminating z and y from Eq. (21), a 4th-degree polynomial equation in x may be obtained, i.e. $p_{123} = 0$. Since three roots of p_{123} necessarily correspond to trivial solutions, the fourth root is real and it may be computed by Vieta's formulas in closed form. The problem admits, thus, a single solution. Of course, it is feasible only if the corresponding cable tensions are nonnegative and the equilibrium is stable.

When the position \mathbf{x} is assigned, \mathbf{r}_1 , \mathbf{r}_2 and \mathbf{r}_3 are unknown. If the Rodrigues parameters e_1 , e_2 and e_3 are chosen to describe the platform orientation, the relations in Eq. (21) assume a particularly favorable structure. Indeed, after letting $\mathbf{r}_i = \mathbf{R}\mathbf{b}_i$, $i = 1 \dots 3$, p_1 , p_2 and p_3 become quartic polynomials in e_1 , e_2 and e_3 . Quartic polynomials in the Rodrigues parameters also emerge from the minors $\det \mathbf{M}'_{1245}$, $\det \mathbf{M}'_{1246}$ and $\det \mathbf{M}'_{1256}$, but they linearly depend on p_1 , p_2 and p_3 , so that they may be discarded. A further quartic emerges by setting $\det \mathbf{M}'_{j456} = 0$ for $j = 1 \dots 3$, so that $(\mathbf{x} + \mathbf{r}_1) [\det \mathbf{M}'_{456,234}] = 0$. The variety defined by the above equation comprises, other than the trivial solution $\mathbf{x} = -\mathbf{r}_1$, also the set of configurations for which

$$p_8 := \det \mathbf{M}'_{456,234} = 0. \quad (22)$$

Equation (22) has, indeed, degree 4 in e_1 , e_2 and e_3 . All minors of \mathbf{M}' not considered so far, namely $\det \mathbf{M}'_{1345}$, $\det \mathbf{M}'_{1346}$, $\det \mathbf{M}'_{1356}$, $\det \mathbf{M}'_{2345}$, $\det \mathbf{M}'_{2346}$ and $\det \mathbf{M}'_{2356}$, yield, instead, sextic equations in the Rodrigues parameters.

It is known that three polynomial equations of the same total degree always admit a Sylvester-type resultant free from extraneous polynomial factors [39]. For the case of three quartics, such a resultant is a 64th-degree polynomial in one of the unknowns. This polynomial may be obtained, in the present case, from p_8 and any two between p_1 , p_2 and p_3 (the resultant of p_1 , p_2 and p_3 is identically nought). However, it is possible to show that the varieties V_{128} , V_{138} and V_{238} only have 48 solutions in common and only 24 of them actually satisfy the aforementioned sextics, thus belonging to V . An effective strategy in order to obtain a univariate 24th-degree polynomial in one of the unknowns consists in deriving a larger set of linearly independent quartics. This may be accomplished by choosing suitable moment poles different from O in the computation of \mathbf{M} , thus deriving additional minors that are linearly independent from the ones emerging from Eqs. (7)-(8). Among them, seven more quartics in the Rodrigues parameters, indeed, exist. By applying Sylvester's dialytic method to the eleven quartics which are altogether available, a univariate 24th-degree polynomial in one of the unknowns, e.g. e_3 , may be finally obtained. Details will be provided in a future paper.

Direct geometrico-static problem

When cable lengths are assigned, the platform configuration has to be determined. Equation (5) provides up to 15 (non-independent) polynomial equations in the platform posture

variables. Among them, Eqs. (21) and (22) are of degree four in e_1 , e_2 and e_3 and degree two in x , y , z .

Equation (1) provides three further relations in the platform posture variables. In particular, one may conveniently consider

$$q_1 := |\mathbf{s}_1|^2 - \rho_1^2 = 0, \quad (23a)$$

$$q_2 := |\mathbf{s}_2|^2 - \rho_2^2 - |\mathbf{s}_1|^2 + \rho_1^2 = 0, \quad (23b)$$

$$q_3 := |\mathbf{s}_3|^2 - \rho_3^2 - |\mathbf{s}_1|^2 + \rho_1^2 = 0, \quad (23c)$$

which, after clearing the denominator $1 + e_1^2 + e_2^2 + e_3^2$, are quadratic in e_1 , e_2 and e_3 . q_1 is also quadratic in the elements of \mathbf{x} , whereas q_2 and q_3 are linear in these variables.

The point-to-point distance relations in Eq. (23) represent the typical constraints governing the forward kinematics of parallel manipulators with telescoping legs connected to the base and the platform by ball-and-socket joints. In particular, the direct position problem of the general Gough-Stewart manipulator depends on six equations of this sort, one of which is equivalent to Eq. (23a) and five more to Eqs. (23b)-(23c). This problem is known to be very difficult and it has attracted the interest of researchers for several years [40, 41]. The DGP of the 33-CDPR appears to be even more complex. In fact, three equations analogous to Eqs. (23b)-(23c), namely *linear* in the components of \mathbf{x} and *quadratic* in the components of Φ , are replaced by relationships that are, at least, *quadratic* in the components of \mathbf{x} and *quartic* in the components of Φ (cf. Eqs. (21) and (22)). The problem appears to be a daunting task and it will be the subject of future research.

5. CONCLUSIONS

This paper studied the kinematics and statics of under-constrained cable-driven parallel robots with less than six cables, in crane configuration. In these robots, kinematics and statics are intrinsically coupled and they must be dealt with simultaneously. This poses major challenges. A geometrico-static model was presented, together with an original and general procedure aimed at effectively solving the inverse and direct position problems. A purely algebraic method, based on a constrained optimization formulation, was provided for the assessment of equilibrium stability. A spatial robot with three cables was considered as a case study, in order to show the feasibility of the presented approach.

REFERENCES

- [1] Ming, A., and Higuchi, T., 1994. "Study on multiple degree-of-freedom positioning mechanism using wires—part 1: Concept, design and control". *Int. Journal of the Japan Society for Precision Engineering*, **28**(2), pp. 131–138.
- [2] Arai, T., and Osumi, H., 1992. "Three wire suspension robot". *Industrial Robot*, **19**(4), pp. 17–22.
- [3] Yang, L. F., and Mikulas, M. M., 1994. "Mechanism synthesis and two-dimensional control designs of an active three-cable crane". *Journal of Spacecraft and Rockets*, **31**(1), pp. 135–144.
- [4] Yamamoto, M., Yanai, N., and Mohri, A., 2004. "Trajectory control of incompletely restrained parallel-wire-suspended mechanism based on inverse dynamics". *IEEE Transactions on Robotics*, **20**(5), pp. 840–850.
- [5] Fattah, A., and Agrawal, S. K., 2006. "On the design of cable-suspended planar parallel robots". *ASME Journal of Mechanical Design*, **127**(5), pp. 1021–1028.

- [6] Heyden, T., and Woernle, C., 2006. "Dynamics and flatness-based control of a kinematically undetermined cable suspension manipulator". *Multibody System Dynamics*, **16**(2), pp. 155–177.
- [7] Michael, N., Kim, S., Fink, J., and Kumar, V., 2009. "Kinematics and statics of cooperative multi-robot aerial manipulation with cables". In Proc. of the ASME 2009 Int. Design Engineering Technical Conferences, pp. paper no. DETC2009–87677.
- [8] Landsberger, S. E., 1984. "Design and construction of a cable-controlled, parallel link manipulator". Master's thesis, Massachusetts Institute of Technology, Dept. of Mechanical Engineering.
- [9] Albus, J., Bostelman, R., and Dagalakis, N., 1993. "The NIST robocrane". *Journal of Robotic Systems*, **10**(5), pp. 709–724.
- [10] Ming, A., and Higuchi, T., 1994. "Study on multiple degree of freedom positioning mechanisms using wires—part 2: Development of a planar completely restrained positioning mechanism". *Int. Journal of the Japan Society for Precision Engineering*, **28**(3), pp. 235–242.
- [11] Kurtz, R., and Hayward, V., 1995. "Dexterity measures with unilateral actuation constraints: the $n + 1$ case". *Advanced Robotics*, **9**(5), pp. 561–577.
- [12] Roberts, R. G., Graham, T., and Lippitt, T., 1998. "On the inverse kinematics, statics, and fault tolerance of cable-suspended robots". *Journal of Robotic Systems*, **15**(10), pp. 581–597.
- [13] Kawamura, S., Kino, H., and Won, C., 2000. "High-speed manipulation by using parallel wire-driven robots". *Robotica*, **18**(1), pp. 13–21.
- [14] Su, Y. X., Duan, B. Y., Nan, R. D., and Peng, B., 2001. "Development of a large parallel-cable manipulator for the feed-supporting system of a next-generation large radio telescope". *Journal of Robotic Systems*, **18**(11), pp. 633–643.
- [15] Kossowski, C., and Notash, L., 2002. "CAT4 (cable actuated truss - 4 degrees of freedom): a novel 4 DOF cable actuated parallel manipulator". *Journal of Robotic Systems*, **19**(12), pp. 605–615.
- [16] Tadokoro, S., Murao, Y., Hiller, M., Murata, R., Kohkawa, H., and Matsushima, T., 2002. "A motion base with 6-DOF by parallel cable drive architecture". *IEEE/ASME Transactions on Mechatronics*, **7**(2), pp. 115–123.
- [17] Behzadipour, S., and Khajepour, A., 2004. "Design of reduced DOF parallel cable-based robots". *Mechanism and Machine Theory*, **39**(10), pp. 1051–1065.
- [18] Pusey, J., Fattah, A., Agrawal, S., and Messina, E., 2004. "Design and workspace analysis of a 6-6 cable-suspended parallel robot". *Mechanism and Machine Theory*, **39**(7), pp. 761–778.
- [19] Merlet, J.-P., 2004. "Analysis of the influence of wires interference on the workspace of wire robots". In On Advances in Robot Kinematics, J. Lenarčič and C. Galletti, eds., Kluwer Academic Publishers, Dordrecht, pp. 211–218.
- [20] Behzadipour, S., and Khajepour, A., 2005. "A new cable-based parallel robot with three degrees of freedom". *Multibody System Dynamics*, **13**(4), pp. 371–383.
- [21] Hiller, M., Fang, S., Mielczarek, S., Verhoeven, R., and Franitz, D., 2005. "Design, analysis and realization of tendon-based parallel manipulators". *Mechanism and Machine Theory*, **40**(4), pp. 429–445.
- [22] Gouttefarde, M., and Gosselin, C. M., 2006. "Analysis of the wrench-closure workspace of planar parallel cable-driven mechanisms". *IEEE Transactions on Robotics*, **22**(3), pp. 434–445.
- [23] Bosscher, P., Riechel, A. T., and Ebert-Uphoff, I., 2006. "Wrench-feasible workspace generation for cable-driven robots". *IEEE Transactions on Robotics*, **22**(5), pp. 890–902.
- [24] Kozak, K., Zhou, Q., and Wang, J., 2006. "Static analysis of cable-driven manipulators with non-negligible cable mass". *IEEE Transactions on Robotics*, **22**(3), pp. 425–433.

- [25] Behzadipour, S., and Khajepour, A., 2006. “Stiffness of cable-based parallel manipulators with application to stability analysis”. *ASME Journal of Mechanical Design*, **128**(1), pp. 303–310.
- [26] Behzadipour, S., and Khajepour, A., 2006. “Erratum: stiffness of cable-based parallel manipulators with application to stability analysis”. *ASME Journal of Mechanical Design*, **128**(11), p. 1366.
- [27] Stump, E., and Kumar, V., 2006. “Workspaces of cable-actuated parallel manipulators”. *ASME Journal of Mechanical Design*, **128**(1), pp. 159–167.
- [28] Gouttefarde, M., Merlet, J.-P., and Daney, D., 2007. “Wrench-feasible workspace of parallel cable-driven mechanisms”. In *Proc. of the 2007 IEEE Int. Conf. on Robotics and Automation*, pp. 1492–1497.
- [29] Bosscher, P., Williams II, R. L., Bryson, L. S., and Castro-Lacouture, D., 2007. “Cable-suspended robotic contour crafting system”. *Automation in Construction*, **17**(1), pp. 45–55.
- [30] Wischnitzer, Y., Shvalb, N., and Shoham, M., 2008. “Wire-driven parallel robot: permitting collisions between wires”. *The Int. Journal of Robotics Research*, **27**(9), pp. 1007–1026.
- [31] Alikhani, A., Behzadipour, S., Vanini, S. A. S., and Alasty, A., 2009. “Workspace analysis of a three DOF cable-driven mechanism”. *ASME Journal of Mechanisms and Robotics*, **1**(4), pp. 041005/1–7.
- [32] Diao, X., and Ma, O., 2009. “Vibration analysis of cable-driven parallel manipulators”. *Multibody System Dynamics*, **21**(4), pp. 347–360.
- [33] Diao, X., and Ma, O., 2009. “Force-closure analysis of 6-dof cable manipulators with seven or more cables”. *Robotica*, **27**(2), pp. 209–215.
- [34] Pham, C. B., Yeo, S. H., Yang, G., and Chen, I.-M., 2009. “Workspace analysis of fully restrained cable-driven manipulators”. *Robotics and Autonomous Systems*, **57**(9), pp. 901–912.
- [35] Bouchard, S., Gosselin, C., and Moore, B., 2010. “On the ability of a cable-driven robot to generate a prescribed set of wrenches”. *ASME Journal of Mechanisms and Robotics*, **2**(1), pp. 011010/1–10.
- [36] Greenwood, D. T., 2006. *Advanced dynamics*. Cambridge University Press, Cambridge.
- [37] Luenberger, D. G., and Ye, Y., 2008. *Linear and nonlinear programming*. Springer, New York.
- [38] Hunt, K. H., 1978. *Kinematic geometry of mechanisms*. Clarendon Press, Oxford.
- [39] Salmon, G., 1885. *Lessons introductory to the modern higher algebra*. Hodges, Figgis, and Co., Dublin.
- [40] Merlet, J.-P., 2006. *Parallel robots*. Springer, Dordrecht.
- [41] Sommese, A. J., and Wampler, C. W., 2005. *The numerical solution of systems of polynomials arising in engineering and science*. World Scientific Publishing, Singapore.

INDICE DEGLI AUTORI

Andrisano Angelo Oreste	165, 259
Bagagli Floriano	337
Bagaria William J.	193
Barbieri Marco	259
Bergamasco Massimo	57
Berselli Giovanni	429
Bertocchi Luca	231
Borghi Corrado	303
Campioni Eleonora	231
Carli Andrea	181
Carminelli Antonio	359
Carricato Marco	443
Castagnetti Davide	283
Catania Giuseppe	3, 359
Cavallari Marco	33
Cenci Stefano	181
Ceschini Lorella	373
Chebbi Ahmed Hachem	127
Cocconcelli Marco	303, 315
Croccolo Dario	411
Dalpiaz Giorgio	33
Dal Re Vincenzo	385
De Agostinis Massimiliano	411
Dettori Andrea	57
Donati Lorenzo	337
Dragoni Eugenio	193, 265, 283, 315
Faretra Marco	165
Ferrari Alberto	303
Foroni Alberto	181
Freddi Alessandro	1, 241
Frisoli Antonio	57
Gherardini Francesco	165
Guerra Alessandro	165
Innocenti Carlo	139
Leali Francesco	165
Mantovani Massimo	181
Marin Tito	155
Medri Gianluca	205
Meneghetti Umberto	97

Montanari Alessio	33
Morri Andrea	373
Mozaffari Fomashi Mohammad	113
Mucchi Emiliano	33
Neumann Manfred	17
Nicoletto Gianni	219
Olmi Giorgio	385
Parenti Castelli Vincenzo	83, 113, 127
Pellicano Francesco	259
Pellicciari Marcello	165
Piccinini Marco	429
Pini Fabio	165
Prati Edzeario	155
Reggiani Barbara	337
Reggiani Giacomo	303
Riello Valerio Giordano	181
Rivola Alessandro	73
Rosati Giulio	181
Rossi Aldo	181
Rubini Riccardo	303, 315
Sancisi Nicola	83
Scirè Mammano Giovanni	265, 315
Segatori Antonio	337
Silvestri Marco	155
Solazzi Massimiliano	57
Sorrentino Silvio	3
Spaggiari Andrea	283
Squatrito Rosario	373
Strozzi Antonio	231
Tasora Alessandro	155
Todaro Ivan	373
Tomesani Luca	337, 373
Troncosi Marco	73, 113
Vassura Gabriele	429
Vergnano Alberto	165
Vertechy Rocco	57
Vincenzi Nicolò	411
Wassermann Johann	17
Zanarini Alessandro	17
Zannoli Diego	83
Zanotti Luca	181